



Universitat Autònoma de Barcelona

MASTER DE INFORMÀTICA AVANZADA

MEMORIA DEL TRABAJO DE INVESTIGACIÓN

ITINERARIO Arquitectura de Computadores y Procesamiento Paralelo

Sistema Adaptativo para la Predicción de Incendios Forestales basado en Estrategias Estadístico-Evolutivas

Autor: Rodríguez Aseretto, R. Darío

Fecha: Barcelona, septiembre 2007

Tutor: Cortés Fité, Ana

Dña Ana Cortés Fité, profesora titular del Departamento de Arquitectura de Computadores y Sistemas Operativos de la Universidad Autónoma de Barcelona

CERTIFICA: que la presente memoria "Sistema Adaptativo para la Predicción de Incendios Forestales basado en Estrategias Estadístico-Evolutivas", ha sido realizada bajo su dirección por R. Darío Rodríguez Aseretto, constituyendo el Trabajo de Investigación del Master de Informática Avanzada Itinerario 1: "Arquitectura de Computadores y Procesamiento Paralelo".

Bellaterra, 3 de septiembre de 2007

Fdo, Ana Cortés Fité

Agradecimientos

Primero que nada, debo agradecer y dedicar este trabajo a mi madre (Eva), padre (Ieka) y mi hermano (Gus). Sin la ayuda de ellos jamás podría haber logrado nada en la vida, gracias por su amor, ayuda y comprensión.

También agradecer a los que cariñosamente llamo “profes” Ana, Tomás y Emilio, porque siempre son linternas que alumbran el camino cuando todo está oscuro. Por supuesto, que no me puedo olvidar de mis compañeros de grupo Mónica y Germán, sinónimos de compañerismo, esfuerzo y trabajo.

A todos los profesores, compañeros y funcionarios del Departamento CAOS, que hacen que el Departamento sea como un hogar. No puedo dejar de nombrar a algunos amigos, que compartimos la misma lucha diaria y que los menciono por orden alfabético: Alex, Alexandre, John y Manuel.

Y por supuesto, a mí querida esposa Diana, porque esto es fruto de nuestros sueños, porque juntos somos fuertes y porque me comería al mundo si está a mi lado. A ella, simplemente, gracias por existir.

Índice general

1. Introducción	1
1.1. Ciencia Computacional	2
1.2. Simulación	3
1.2.1. Tipos de Simulación	4
1.2.2. Simuladores de Propagación de Incendios Forestales	5
1.3. Cómputo Paralelo y Distribuido	7
1.3.1. Paradigma de programación	8
1.4. Objetivos	10
1.5. Organización del trabajo	11
2. Métodos de Predicción de Incendios Forestales	12
2.1. Método Clásico	14
2.2. Métodos Guiados por los Datos	16
2.2.1. Solución Única	17
2.2.2. Solución Múltiple	18
3. Métodos Guiados por los Datos	21
3.1. Método Evolutivo basado en Algoritmos Genéticos	22
3.1.1. Descripción del Método	23
3.1.2. Implementación del Método	24
3.1.3. Ventajas	28
3.1.4. Desventajas	29
3.2. Método Estadístico - S^2F^2M (Statistical System for Forest Fire Management)	30
3.2.1. Descripción del Método	31
3.2.2. Implementación del Método	33
3.2.3. Ventajas	38
3.2.4. Desventajas	38
4. $SAPIFE^3$ - Sistema Adaptativo para la Predicción de Incen- dios Forestales basado en Estrategias Estadístico-Evolutivas	40
4.1. Bases Teóricas y Prácticas de la Propuesta	41
4.2. Propuesta - Intregación del Método Estadístico con el Evolutivo	42

4.3.	Descripción del Método	45
4.4.	Implementación del Método	47
4.5.	Implementación en un Esquema Paralelo Master/Worker . . .	48
5.	Resultados Experimentales	51
5.1.	Mapas Sintéticos	53
5.1.1.	Sintético I	54
5.1.2.	Sintético II	58
5.2.	Mapas Reales	62
5.2.1.	Plot 520	63
5.2.2.	Plot 751	66
6.	Conclusiones y Trabajo Futuro	70
6.1.	Consideraciones	70
6.2.	Conclusiones	71
6.3.	Trabajo Futuro	74

Índice de figuras

1.1. Ciencia Computacional	3
1.2. Modelos de Propagación	6
1.3. Estructura básica de un simulador de fuego	7
1.4. Modelo SPMD (a) Modelo Pipeline (b) Modelo Divide/Conquer (c) Modelo Master/Worker (d)	8
2.1. Estructura de pipeline del FireLib	13
2.2. Estructura del simulador fireSim	14
2.3. Predicción Clásica	15
2.4. Predicción Guiada por los Datos	17
2.5. Solución Múltiple	19
3.1. Taxonomía de los Métodos de Predicción de Incendios Forestales	21
3.2. Metodo Evolutivo de Predicción de Incendios Forestales . . .	23
3.3. Diagrama de flujo convencional de un algoritmo genético . . .	24
3.4. Individuo de la población	25
3.5. Ej. Población de 10 individuos	26
3.6. Cálculo de Error	26
3.7. Comportamiento del error según el número de iteraciones . .	27
3.8. Cruce y Mutación	28
3.9. Método Estadístico de Predicción de Incendios Forestales . .	32
3.10. Ej. Estructura de S^2F^2M	33
3.11. Ej. Escenarios de simulación	34
3.12. Archivos de parámetros de S^2F^2M	34
3.13. Ejemplo de Matriz de Probabilidades	35
3.14. Esquema del método S^2F^2M	37
3.15. Resultado de baja calidad	39
4.1. Integración de las simulaciones de los distintos métodos . . .	44
4.2. Estructura de $SAPIFE^3$	45
4.3. Vista en tres dimensiones de la salida de $SAPIFE^3$	46
4.4. Esquema de Implementación de $SAPIFE^3$	47
4.5. Estructura parbegin-parend de $SAPIFE^3$	49

4.6. Pseudocódigo Simple de la Implementación de <i>SAPIFE</i> ³ en el modelo Master/Worker	49
5.1. Función Fitness	52
5.2. Sintético I	54
5.3. Mapa sintético I, errores medidos en la etapa de ajuste y predicción	55
5.4. Mapa sintético I, comparativa con los tres métodos	57
5.5. Primer Periodo del Sintético II	58
5.6. Segundo Periodo del Sintético II	59
5.7. Mapa Sintético II, errores medidos en las etapas de ajuste y predicción	60
5.8. Mapa Sintético II, comparativa entre los tres métodos	61
5.9. Impacto de los cambios bruscos en los tres métodos	62
5.10. Vista del área de Gestosa, donde se realizaron las quemas prescriptas	63
5.11. Mapa de fuego real, plot 520	64
5.12. Plot 520, errores obtenidos en la etapa de ajuste	65
5.13. Plot 520, errores obtenidos en la etapa de predicción	66
5.14. Mapa de fuego real, plot 751	67
5.15. Plot 751, errores obtenidos en la etapa de ajuste	68
5.16. Plot 751, errores obtenidos en la etapa de predicción	69

Capítulo 1

Introducción

La primera parte del siglo XXI ha supuesto un patrón inusual en términos de frecuencia, severidad y áreas quemadas en números muy elevados en diversas partes del globo, afectando muy seriamente a la seguridad humana. En años anteriores los fuegos han provocado pérdidas devastadoras; incendios severos en Portugal, Francia y España en el verano, seguido por los grandes incendios en EEUU-California. Las experiencias compartidas de estos fuegos han acentuado la importancia de aprovechar al máximo los recursos e investigaciones disponibles para encontrar mejores maneras de combatir incendios.

Para reducir al mínimo la amenaza de pérdidas, los encargados de combatir el fuego deben planear estrategias de protección apropiadas. Un requisito previo para este planeamiento es la capacidad de determinar cual será el comportamiento del incendio, usando esta información, los encargados pueden establecer las prioridades para que las actividades de combate reduzcan el peligro del mismo. Para ello, es necesario el uso de un simulador basado en un modelo matemático, el cuál debe predecir por dónde el incendio se va a propagar y en qué momento, pero esto necesita un gran número de parámetros de entrada que casi siempre es muy difícil de obtener.

En las ciencias ambientales se ha ido incrementando el uso de modelos para efectuar predicciones. Pero los modelos ofrecen una serie de limitaciones debido a la falta de precisión o de una incorrecta validación. Algunas de estas restricciones pueden producir consecuencias dramáticas. Además, es común que tales limitaciones sean impuestas por los propios modelos, que para operar suelen precisar datos en un gran número de variables de entrada, y dichas variables, por las propias características del modelo matemático, suelen exhibir algún grado de incertidumbre. Asimismo, las soluciones numéricas solo son una aproximación a la realidad, aún sin considerar las traducciones de estas soluciones cuando son llevadas a cabo mediante ordenadores. Y por ende, la validación de modelos, así como la verificación y calibración se torna una tarea árdua. [1].

Un simulador que proporciona una buena predicción es la base de cualquier

sistema para soporte a la gestión de recursos en el combate contra incendios forestales. Considerando los simuladores de fuego ya existentes, el método propuesto debe ser implementado para trabajar con cualquiera de éstos sin modificar su funcionamiento interno, por lo tanto, la propuesta del nuevo método solo opera sobre los parámetros de entrada a través de un análisis de la salida que brinda dichos simuladores.

Para llevar a cabo esta tarea, es necesario un gran poder de cómputo con restricciones de tiempo real, por lo que también se han analizado técnicas y alternativas para implementar el método. La complejidad del problema requiere de la Ciencia Computacional para coordinar la distintas áreas involucradas dentro de la investigación (Física, Matemática, Ingeniería e Informática).

Según Fracis Bacon, el paradigma científico, se define como el proceso de formar hipótesis y verificarlas a través de experimentos; hipótesis con éxito son vistas como modelos que explican fenómenos en el mundo [2]; partiendo de esta premisa se elaboró una serie de experimentos prácticos para contrastarlos con casos reales con el objetivo de comprobar la efectividad y viabilidad del método.

1.1. Ciencia Computacional

Muchos de los descubrimientos realizados por los científicos e ingenieros no hubieran sido posibles sin las computadoras, éstas han revolucionado la manera en que los científicos realizan su trabajo y se podría decir que las computadoras son el laboratorio de la ciencia moderna.

Tradicionalmente, la ciencia fue hecha en un laboratorio como combinación de la teoría y de la experimentación física (con cálculos a mano), pero las computadoras han hecho posible una nueva manera de hacer ciencia y de gran alcance con simulación numérica. La simulación numérica es el proceso de modelar matemáticamente un fenómeno físico, y después hacerlo funcionar en un experimento con el modelo matemático. La “Ciencia Computacional” desempeña un papel importante en esta nueva manera de hacer ciencia, de crear, de evaluar, y de refinar los modelos matemáticos usados para simular los fenómenos físicos. [3].

La ciencia computacional no debe ser confundida con la Informática (Computer Science). La primera se enfoca en problemas científicos o de ingeniería, y junto con la matemática y la informática obtiene un entendimiento perfeccionado de los problemas. Por otra parte, la informática centra su interés en el computador mismo (estudio matemático del cómputo, computadores y el procesamiento de la información). Aunque las áreas son bastante distintas, muchos de los temas típicamente considerados pertenecientes al dominio de “computer science” son de mucho valor en la ciencia computa-

cional. [4].

Entonces, a partir de estos antecedentes Tapia y Lanius [3] definen la “Ciencia Computacional” como: La ciencia de cómputo es un campo interdisciplinario en la intersección de tres dominios: de las matemáticas, de la informática, y de las ciencias biológicas y físicas. Los científicos de esta área utilizan las herramientas de la informática y de las matemáticas para estudiar problemas de la ciencia, la sociología, la física, ingeniería, etc. La mayor parte de los problemas en los cuales los científicos trabajan implican grandes cantidades de datos y una gran cantidad de variables. Con los avances en informática y métodos numéricos, los matemáticos y los científicos pueden trabajar juntos, modelar y solucionar los problemas que eran imposibles tratar hace por lo menos diez años atrás.

En esta clase de ciencia se hace más que usar una computadora para encontrar soluciones a problemas de modelos matemáticos desarrollados por científicos de otras áreas. También desarrollan nuevas herramientas matemáticas, teorías y nuevos métodos numéricos que mejoran la exactitud y la velocidad de métodos existentes. El problema de la simulación de incendios forestales encaja perfectamente en esta forma de hacer ciencia como puede observarse en la figura 1.1.

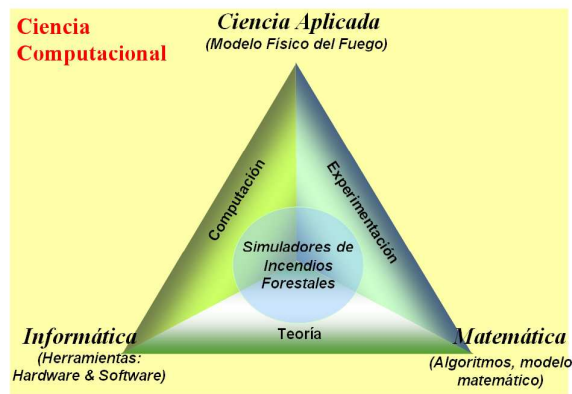


Figura 1.1: Ciencia Computacional

1.2. Simulación

La simulación es una técnica numérica para conducir experimentos en una computadora digital. Estos experimentos comprenden ciertos tipos de relaciones matemáticas y lógicas, los cuales son necesarios para describir el comportamiento y la estructura de sistemas complejos del mundo real a través de largos periodos de tiempo. [5]

El fundamento racional para usar la simulación en cualquier disciplina, no solo en la predicción de incendios forestales, es la necesidad del hombre de adquirir algún tipo de conocimiento con respecto a la predicción del futuro. Esto es tan antiguo como la historia de la humanidad; antes de siglo XVII, esa indagación estaba casi limitada a métodos puramente deductivos de los filósofos como Platón, Aristóteles, Euclides y otros, denominada por algunos críticos como Hans Reichenbach “filosofía especulativa” que pretendió dar un conocimiento de los principios más generales que gobiernan al universo.

Se debe considerar que estos sólo eran simples intentos de dar una aproximación a lo que hoy conocemos como explicación científica a través de simples analogías con las experiencias de la vida cotidiana. [6].

Pero hay que dejar bien claro, que ninguna parte sustancial del universo es tan simple como para comprenderse y controlarse sin abstracción. La abstracción consiste en reemplazar la parte del universo bajo consideración, por un modelo de estructura similar pero más simple. Entonces, podemos definir el modelo científico como una abstracción de algún sistema real. Y estos deben ser simples y realistas. Desafortunadamente los modelos realistas raras veces son simples y los simples raras veces son realistas.

Es de conocimiento cuan difícil es realizar un modelo que represente de manera realista un suceso físico y de las dificultades y limitaciones del modelo matemático, pero esto, aún debe ser codificado en un lenguaje de programación y dicha traducción, del modelo matemático al computador, deja aún en mayor evidencia los resultados dados en la simulación por computadora.

1.2.1. Tipos de Simulación

Los tipos de simulación que se pueden efectuar se clasifican de la siguiente forma [7]:

- Simulación Discreta: modelación de un sistema por medio de una representación en la cual el estado de las variables cambian en instantes de tiempos separados. En términos matemáticos el sistema sólo puede cambiar en instantes de tiempos contables.
- Simulación Continua: modelación de un sistema por medio de una representación en la cual las variables de estado cambian continuamente en el tiempo. Típicamente, los modelos de simulación continua involucran ecuaciones diferenciales que determinan las relaciones de las tasas de cambios de las variables de estado en el tiempo.
- Simulación Combinada Discreta-Continua: modelación de un sistema por medio de una representación en la cual unas variables de estado cambian continuamente con respecto al tiempo y otras cambian en instantes de tiempos separados.

- Simulación Determinística y/o Estocástica: una simulación determinística es aquella que utiliza únicamente datos de entrada determinísticos, no utiliza ningún dato de entrada azaroso. En cambio, un modelo de simulación estocástico incorpora algunos datos de entrada azarosos al utilizar distribuciones de probabilidad.
- Simulación Estática y Dinámica: La simulación estática es aquella en la cual el tiempo no juega un papel importante, en contraste con la dinámica en la que si es muy importante.
- Simulación con Orientación hacia los Eventos: modelaje con un enfoque hacia los eventos, en el cual la lógica del modelo gira alrededor de los eventos que ocurren instante a instante, registrando el estado de todos los eventos, entidades, atributos y variables del modelo en todo momento.
- Simulación con Orientación hacia Procesos: modelaje con un enfoque de procesos, en el cual la lógica del modelo gira alrededor de los procesos que deben seguir las entidades. En cierta forma, es un modelaje basado en un esquema de flujograma de procesos, en el que se hace un seguimiento a la entidad a través de la secuencia de procesos que debe seguir.

A partir de estas clasificaciones, se concluye que la mayoría de los simuladores de incendios forestales realizan el tipo de Simulación Determinística y/o Estocástica partiendo de la base causa-consecuencia. Por lo tanto, una serie de condiciones geográficas, climáticas y ambientales, traducidas como variables de entrada determinan el comportamiento del fuego dentro del modelo y el gran problema es que dichas variables son exógenas al modelo con una gran dificultad para medirlas, además de modificarse dinámicamente durante el transcurso del incendio.

1.2.2. Simuladores de Propagación de Incendios Forestales

Los sistemas de simulación de incendios operan a distintas escalas de tamaños para construir el mapa de propagación y se dividen en procesos locales y globales. Los simuladores cuando trabajan a escala global consideran al incendio como un todo o bien una unidad geométrica para representar la línea de fuego en el espacio y tiempo. Asimismo, cuando operan a escala local utilizan puntos o celdas para construir la línea de fuego. La forma de calcular la propagación a nivel local puede ser de manera distinta como muestra la figura 1.2. En el modelo de propagación elíptico, para cada punto se construye una elipse calculando de esta manera la excentricidad de la misma, para determinar el sentido de la propagación para el siguiente punto. En el modelo de propagación por celdas, se puede ver como cada celda determina

en qué momento la celda vecina es alcanzada por el fuego. Una vez procesado todo el modelo local el sistema integra todos los procesos para construir el nuevo frente de fuego [8] [9] (Ver figura 1.2).

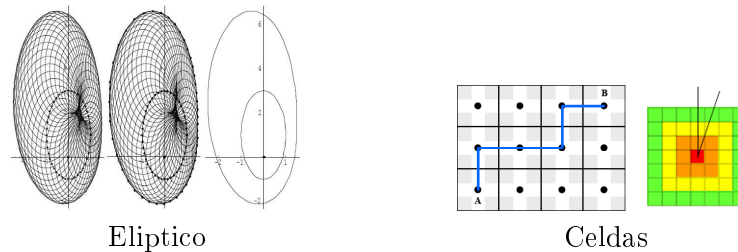


Figura 1.2: Modelos de Propagación

Los simuladores de incendios deben considerar tres tipos de parámetros de entradas:

- Los Iniciales, donde se especifican las coordenadas de dónde se inició el incendio o bien cuál es la situación actual del mismo.
- Los Estáticos, son aquellos que no varían en el tiempo y generalmente están almacenados en bases de datos como los concernientes a la topografía del terreno y la vegetación.
- Los Dinámicos, que varían en el transcurso del tiempo y son más difíciles de medir.

Por lo tanto, todos estos tipos de parámetros deben ser procesados en los modelos de propagación dentro del simulador, para luego dar una salida indicando un nuevo mapa de la línea de fuego. La figura 1.3 muestra cuál sería la estructura básica de un simulador de incendios forestales, donde los tres tipos de parámetros comentados se reflejan en la columna denominada Tipo. Las tres columnas restantes (Entrada, Modelos de Propagación y Salida) reflejan la forma en que el simulador trabaja y se describen a continuación:

- Parámetros de entrada: aquí se ingresa toda la información que concierne al ambiente físico y ambiental, en el cual se desarrolla el incendio, el lugar donde se inició el fuego (línea de fuego inicial), el tipo de vegetación, la topografía del terreno, las condiciones climatológicas, etc.
- Modelos de Propagación: El simulador procesa la información aplicando los parámetros a cada entidad local para luego integrarlos en un modelo global de propagación.
- Salida: El simulador construye el nuevo frente de fuego para el instante de tiempo $t + \Delta$.

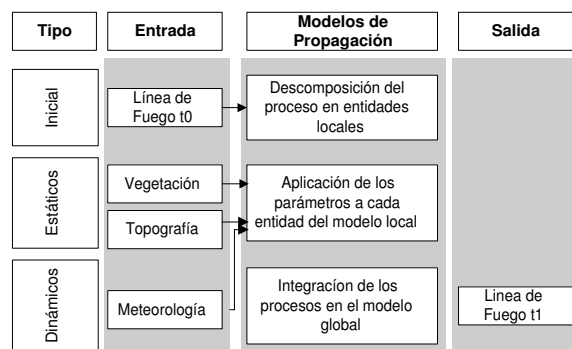


Figura 1.3: Estructura básica de un simulador de fuego

Algunos de los simuladores de fuego más destacados y de mayor uso en la actualidad, están basados en modelos de propagación de Rothermel [10], que es un modelo matemático semi-empírico. El citado modelo esta basado en correlaciones empíricas, obtenidas a partir de datos experimentales, y aunque los resultados de este tipo de modelos son fiables, tienen un rango limitado dentro de las condiciones específicas en que fueron efectuados los experimentos.

1.3. Cómputo Paralelo y Distribuído

La élite de la super computación se encuentra avocada a resolver problemas complejos y de gran magnitud, en cuanto a datos y cómputos, como los sistemas que tratan yacimientos petroleros, la atmósfera, los campos electromagnéticos, los océanos, los metalúrgicos, el aparato circulatorio de los seres humanos, la corteza terrestre, los suelos y las cimentaciones, así como muchos sistemas ambientales imposibles de resolver en un tiempo razonable por un computador o aplicación en serie. Para afrontar estos desafíos se requiere de lo último en recursos computacionales.

La computación en paralelo es una técnica que nos permite distribuir esa gran carga computacional entre muchos procesadores. Como es de conocimiento, una de las mayores dificultades del procesamiento en paralelo es la coordinación de las actividades de los diferentes procesadores y el intercambio de información entre los mismos. Los métodos de descomposición del dominio introducen desde la formulación matemática del problema una separación natural de las tareas y simplifican considerablemente la transmisión de información. En ellos, los sistemas físicos son descompuestos en varios subdominios, frecuentemente muchos, lo cual se refleja en la ingeniería de software del código correspondiente. Así, mediante los métodos de descomposición de

dominio, la programación orientada a objetos y esquemas de paralelización que usan el paso de mensajes, es posible construir aplicaciones interesantes y útiles, relevantes para la solución de muchos problemas de ciencia e ingeniería entre los cuales se encuentra la predicción de incendios forestales.

Existe una clasificación de cuales son las plataformas de arquitectura paralela, pero lo que verdaderamente nos interesa para desarrollar este trabajo son los paradigmas de programación para poder construir nuestra aplicación. Teniendo en cuenta que los métodos debieron ser probados en un cluster de computadoras, a continuación se presenta un resumen de las posibilidades que se estudiaron para desarrollar todos los métodos en este trabajo.

1.3.1. Paradigma de programación

Para afrontar el desafío de disminuir el tiempo de simulación es necesario comprender cuáles son los modelos de algoritmos propuestos para paralelizar nuestra aplicación, sabemos que distribuir la carga mejora la prestaciones y podemos obtener mejores resultados, tenemos un número importante de simulaciones con restricciones de tiempo real para cada paso del método, es por ello que se tuvo en cuenta las diferentes alternativas que caen dentro del paradigma de programación paralela.

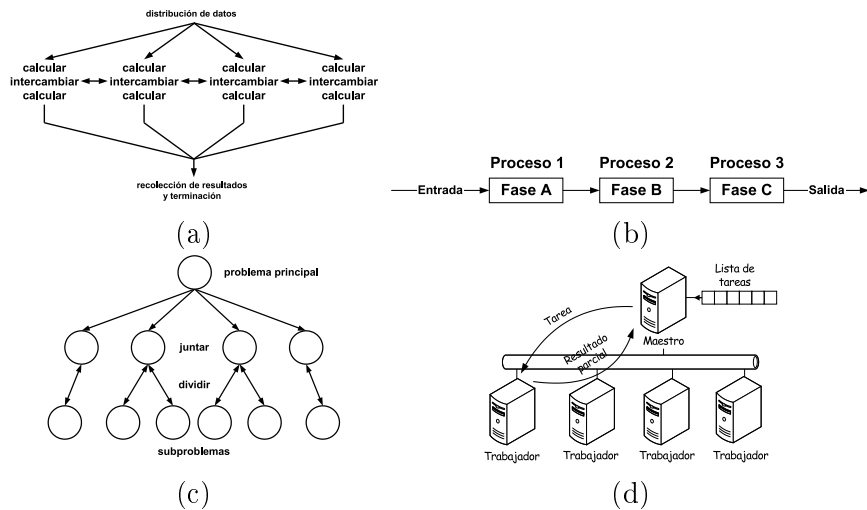


Figura 1.4: Modelo SPMD (a) Modelo Pipeline (b) Modelo Divide/Conquer (c) Modelo Master/Worker (d)

- SPMD: El paradigma SPMD (Single Program Multiple Data) es el más usado de todos los paradigmas para estructurar un programa parale-

lo. Cada proceso ejecuta el mismo código sobre una porción de datos diferente. El código envuelve la distribución de los datos entre los procesadores disponibles. La comunicación entre los procesadores vecinos es proporcional al tamaño de los límites de los elementos, mientras que el cómputo es proporcional al volumen de los elementos. A veces, después del intercambio de información, puede ser necesaria una sincronización global entre todos los procesos. (figura 1.4.a).

- Pipeline: En el paradigma pipeline, el paralelismo está basado en la descomposición funcional. Se identifican las tareas del algoritmo, que son capaces de operar concurrentemente y cada proceso ejecuta una parte pequeña del algoritmo general. Cada proceso corresponde a una etapa del pipeline y es responsable de una tarea particular. Los patrones de comunicación son muy simples y la comunicación puede ser completamente asincrónica. Este método se utiliza habitualmente en aplicaciones de reducción de datos o de procesamiento de imágenes. (figura 1.4.b).
- Divide/Conquer: El paradigma divide y vencerás se usa en el desarrollo de algoritmos secuenciales. Un problema es dividido en dos o más subproblemas. Cada uno de estos subproblemas es resuelto independientemente y se combinan sus resultados para dar el resultado final. Generalmente los problemas menores son sólo instancias menores del problema original. El proceso de partición y de recombinación hace uso del paralelismo, pero estas operaciones requieren comunicación. En cambio, los subproblemas son independientes, sin que resulte necesaria la comunicación entre ellos. En resumen, en este paradigma se realizan tres operaciones: dividir, computar y juntar. La aplicación se organiza en un árbol virtual siendo los procesos hojas los encargados de ejecutar las tareas. (figura 1.4.c).
- Master/Worker: Este paradigma consiste en dos tipos de elementos básicos, un master y varios workers. El master es responsable de dividir el problema en tareas menores, distribuir las mismas entre los procesos workers y recibir los resultados parciales de los workers, para producir finalmente el resultado global. Los procesos workers ejecutan un ciclo simple: obtener un mensaje con la tarea, procesar la tarea y enviar el resultado al proceso master. La comunicación sólo se produce entre el master y los workers. (figura 1.4.d).

El modelo Master/Worker es el que mejor se adapta a las necesidades de los nuevos métodos propuestos, porque todos los modelos tratados en este trabajo, como S^2F^2M y el Evolutivo, utilizan miles de simulaciones con distintos

conjuntos de parámetros para elaborar las predicciones. Más adelante, en el capítulo 3 se describen con mayor detalle como funcionan estos métodos y su implementación. De modo general, se puede decir que cada simulación es enviada a un Worker por medio de un Master que es el encargado de procesar la información global del sistema y recopilar todos los resultados parciales de manera a elaborar, ya sea una población de individuos o un mapa probabilístico de la propagación del fuego dependiendo del método utilizado.

1.4. Objetivos

Los objetivos propuestos al iniciar este trabajo son los siguientes:

- Mejorar la calidad de la predicción: El principal problema que presenta la simulación de incendios forestales son las diferencias entre el mapa simulado y el incendio real, esto produce consecuencias catastróficas en una situación real en donde todas las estrategias y planificaciones de la lucha contra el incendio puede costar vidas humanas. Una buena calidad en la predicción si bien no representa de manera exacta, aunque esto sería lo ideal, la propagación de fuego debe moverse en un rango de error aceptable de tal manera a que sirva para orientar a los responsables, de los riesgos del incendio. Consecuentemente, los métodos propuestos han sido muy útiles a la hora de evaluar el comportamiento del incendio pero aún deben ser mejorados y es el punto en el cual se enfoca la propuesta de este trabajo.
- Reducir el tiempo de Simulación: Otro problema, no menos importante que el anterior, es la duración de la simulación, y aquí se comienza a mencionar la restricción de *Tiempo Real*, una información desfasada en el tiempo es una información inútil. Lo que se busca describir con esto es que para elaborar una predicción del comportamiento de incendio se debe dar una respuesta antes que el fuego real se haya propagado por la zona que el simulador (en el caso que el mapa simulado sea semejante al del fuego real) lo haya predecido, todo esto en un tiempo razonable para que se puedan tomar las medidas correspondientes.

Los métodos propuestos hasta el momento, han logrado una calidad de predicción razonable que han mejorado los métodos tradicionales de predicción, pero aún presentan algunos problemas, los cuales serán abordados en este trabajo. Una de las limitaciones de uno de los métodos, como S^2F^2M , es la cantidad excesiva de tiempo que necesita para dar una buena predicción y unas de las limitaciones del Evolutivo es la adaptación a los cambios de las condiciones ambientales durante el transcurso del incendio. La propuesta es

analizar las fortalezas de los dos métodos, y conociendo su funcionamiento integrarlo en una única metodología para obtener mejor calidad de predicción en un menor tiempo de simulación.

1.5. Organización del trabajo

El Capítulo 2 aborda todo lo referente a los métodos utilizados para la predicción de incendios forestales, como el método Clásico de predicción, en donde cita una breve descripción de los simuladores de incendios más utilizados, también se discuten los métodos en los cuales se aplican la metodología guiada por los datos en sus dos vertientes: solución única y solución múltiple, para luego converger en el tipo de método propuesto en este trabajo.

El Capítulo 3 trata todo lo relacionado a los métodos Evolutivo y Estadístico de predicción de incendios forestales, con una descripción del funcionamiento, también se explica toda la terminología que se utiliza en ambos métodos y la manera que fueron implementados. Además, se describen las ventajas y desventajas y su influencia en el resultado final.

El Capítulo 4 expone la propuesta de combinar los métodos, analizando los aspectos comunes entre ambos, los que sirvieron de base para la combinación, además se mencionan las ventajas que se pretenden alcanzar y la forma de implementación en un esquema paralelo. Igualmente, se describen las modificaciones que se tuvieron que realizar a los métodos utilizados para construir el método adaptativo de predicción.

En el Capítulo 5 se detalla la descripción y el tipo de la experimentación utilizado, las que se dividen en casos reales y sintéticos, donde se evalúan los resultados de los distintos métodos, exponiendo las ganancias y pérdidas, estableciendo gráficas comparativas para demostrar la viabilidad de la propuesta de este trabajo.

El Capítulo 6 expone las conclusiones alcanzadas en el presente trabajo, como así también propone algunas líneas abiertas que podrían ser objeto de futuros trabajos de investigación.

Capítulo 2

Métodos de Predicción de Incendios Forestales

Todos los métodos para la predicción de incendios forestales cuentan con una herramienta principal, un simulador de fuego y la mayoría de ellos adoptan uno de los tres diversos acercamientos enumerados a continuación [12]:

Propagación del frente de fuego: Donde la tasa de propagación hasta el frente de fuego, se calcula a partir de las condiciones ambientales locales, usando el principio de Huygens de propagación de onda y se proyecta adelante en el tiempo, ejemplos de simuladores que utilizan este modelo: SIROFIRE [13], FARSITE [14], Prometheus [15]. La simulación se ejecuta en un ordenador personal, generalmente usa una precisión GIS¹ para representar las características del paisaje, pero esto requiere generalmente algoritmos complejos para manejar la geometría de los frentes convergentes del fuego, por ejemplo, de Knight y de Coleman [16]. Estos tipos de simuladores siguen solamente el frente del fuego y se requiere un procesamiento posterior para estimar la intensidad del fuego o el período ardiente en cualquier localización.

Basado en Física: Son simuladores que calculan la interacción del fuego con la atmósfera, modelan el flujo del fluido calentado (mecánica de fluidos), de la radiación, de la combustión, etc., solucionando las ecuaciones diferenciales relevantes sobre una malla fina en el espacio y el tiempo, algunos ejemplos son: Clark [17], Linn [18], Serón [19]. Estos simuladores requieren recursos de cómputo significativos, (es decir superordenadores), debido al cómputo detallado del sistema entero. Estos simuladores no son convenientes para el uso operacional con las capacidades de cómputo actuales. También, la alta fidelidad que exige la física es desproporcionada con la calidad de datos del combustible y meteorológicos con que se cuenta como parámetros de entrada.

¹Geographic Information Systems (GIS): Son sistemas informáticos capaces de montar, de almacenar, de manipular, y de exhibir la información con referencia geográfica.

Autómatas Celulares: Dos ejemplos de este tipo de metodología son: Berjak [20] y Dunn [21]. En estos casos, mayormente simula la tasa de tiempo para la propagación indicando el tiempo por medio del retardo que sufre el frente de fuego para alcanzar cada celda. Estos simuladores simplemente indican en qué tiempo fue alcanzado por el fuego cada celda, los algoritmos son mucho más simples, pero la geometría de la rejillas pueden influenciar en la forma del fuego.

Todos los métodos descritos más adelante cuentan con un simulador en común, propuesto por Colins D. Bevins, llamado fireSim que implementa la librería denominada firelib². El simulador utiliza una aproximación basada en celdas, en donde el terreno es dividido en mallas y por medio de la relación entre los vecinos colindantes se evalúa si la celda ha sido alcanzada por el fuego y en qué instante.

El fireSim es del tipo de simulador paramétrico determinístico y de eventos discretos, implementado como un simulador autómatas celular, en donde el proceso es por medio de una estructura de tipo pipeline dividido en cuatro etapas (ver figura 2.1). A pesar de que las características del combustible varían a lo largo del terreno, éstas suelen ser consideradas invariantes dentro del marco de tiempo de una simulación.

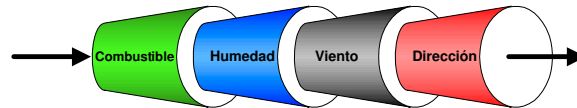


Figura 2.1: Estructura de pipeline del FireLib

Las funciones que trabajan en las distintas etapas del pipeline son [1]:

- FireFuelCombustion: En esta función se procesa el catálogo de combustible que cuenta con varios modelos clasificados por el tipo de partículas de combustibles (hierbas, hojas, ramas, leña menuda, etc).
- FireFireSpreadNoWindNoSlope: Debido a que el modelo de Rothermel [10] usa la humedad del combustible para determinar el término de disipación de calor en la intensidad del fuego, necesita un cálculo previo sin viento ni inclinación, para despejar la variable por medio de esta función.

²fireLib [23] es una librería que encapsula el algoritmo de comportamiento de fuego BEHAVE [24], pero está modificada para aplicaciones interactivas que trabajan con simulaciones de crecimientos basados en celdas.

- FireFireSpreadWindSlopeMax: En esta función se calcula el máximo índice de propagación y la dirección de dicho máximo
- FireFireSpreadAtAzimuth: En esta función se determina el comportamiento del fuego en cualquier dirección utilizando el modelo de crecimiento elíptico desarrollado por Anderson [22] y empleado en BEHAVE [24].

La estructura del simulador fireSim se divide en 5 bloques como se puede apreciar en la figura 2.2, mereciéndose una mención especial el bucle principal, que opera a través de un algoritmo de *contagio* aplicado a las 8 vecinas de una celda analizada en un momento determinado. Por cada celda quemada, el algoritmo analiza las ocho celdas vecinas y determina el tiempo de ignición de las mismas, siempre y cuando las condiciones (humedad, viento, etc.) permitan que el fuego llegue hasta ellas.

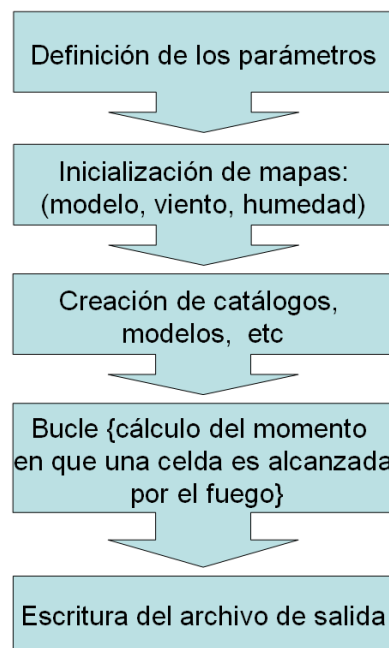


Figura 2.2: Estructura del simulador fireSim

2.1. Método Clásico

Actualmente, la predicción de propagación de incendios forestales que utiliza el método clásico, incluye el uso de un simulador pudiendo ser cualquiera de

los tipos descritos anteriormente. Dependiendo del tipo de simulador los resultados pueden ser más o menos precisos con respecto a la similitud con la propagación real, pero todos éstos son dependientes de la calidad o de la estimación correcta de los valores de los datos de entrada.

Un simulador, primeramente, debe ser inicializado con los datos respecto a las coordenadas de inicio del fuego y los parámetros que describen el entorno del incendio (vegetación, condiciones meteorológicas, etc.) en un instante de tiempo, y luego el simulador brinda una salida indicando cuál será el nuevo frente de fuego para un instante posterior, tal como muestra la figura 2.3.

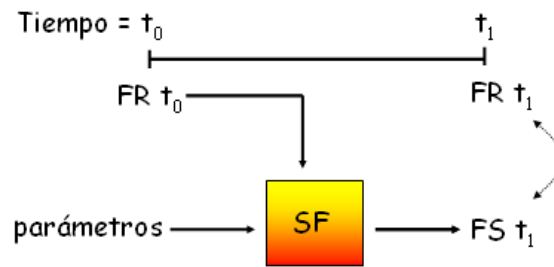


Figura 2.3: Predicción Clásica

Si bien, la simulación en este punto no representa un problema en cuanto a la restricción del tiempo, sí presenta un serio problema de diferencias con el fuego real. Pero, como se mencionó anteriormente la causas de los problemas del modelo es la imprecisión de los parámetros de entrada generalmente por una mala lectura de los mismos. Además, aquellos parámetros que no pueden ser leídos o medidos deben ser estimados lo que provocan serias limitaciones para la predicción del incendio.

A todo esto, se le suma el comportamiento dinámico de las variables medio-ambientales, en donde la lectura de dichas variables se deben efectuar continuamente durante el transcurso del incendio para retroalimentar el simulador, poco probable por la dificultad que implica dicho proceso.

Entonces, el simulador no es la causa verdadera de una mala predicción, haciendo una suposición en donde se pueda medir todas los parámetros de entradas, además de medir la variables dinámicas, y suponiendo que todas sean correctas para introducir dentro del simulador, este podría reproducir una salida buena de la propagación del fuego comparándose con el fuego real (*FR*). Por lo tanto, el verdadero reto se encuentra en conseguir los parámetros óptimos para un incendio que se quiere predecir y he aquí donde nacen las propuestas de métodos que son guiados por los datos (*S²F²M* y Evolutivo), que se describen en el capítulo siguiente.

2.2. Métodos Guiados por los Datos

La propuesta de estos métodos se basa en la búsqueda de estrategias para poder lidiar con las limitaciones descritas en la sección anterior y se considera al simulador de fuego como una herramienta más dentro del modelo (caja negra). Para ello, se cuenta con un método probabilístico que calcula la probabilidad de ignición para una porción del terreno y un método evolutivo que tiene por objetivo la búsqueda del mejor conjunto de parámetros para un incendio con ciertas condiciones atmosféricas y ambientales. Ambos métodos trabajan de formas distintas, pero como se explica más adelante tienen aspectos comunes, por lo que se propone la combinación de los mismos aprovechando de esta manera lo mejor de cada uno, a fin de obtener mejores resultados.

A diferencia del método clásico de predicción, en el método propuesto se introduce una etapa de ajuste, antes de la etapa de predicción. Esta nueva etapa, corresponde a incluir algún método basado en cómputo intensivo [25] que dependiendo de la metodología, se centre en obtener el conjunto de parámetros que mejor prediga el comportamiento del fuego o un valor probabilístico sobre la matriz que representa el terreno en ignición. Dicho conjunto de parámetros o valor de probabilidad será utilizado en la etapa de predicción para un instante de tiempo posterior.

Básicamente, los métodos guiados por datos consisten principalmente en proteger el método clásico de predicción e incluir previamente una etapa de ajuste. Entonces, a partir de un incendio, se efectúa un punto de control, en dicho punto se toma el mapa de la propagación del fuego desde el instante t_0 hasta el instante t_1 , tal como muestra la figura 2.4 y por otro lado, se efectúa simulaciones con distintos conjuntos de parámetros buscando los resultados que por medio de una comparativa consigan el conjunto/s de parámetros que mejor reproduzcan el comportamiento del fuego hasta dicho instante. Una vez que se obtienen los parámetros que mejor representen la propagación del incendio se efectúa la predicción para un instante posterior t_2 .

Este nuevo paradigma que va más allá de las técnicas tradicionales de interpretar los datos, convierte a las simulaciones y las medidas de control en un sistema simbiótico de regeneración continua de gran alcance. La capacidad de cambiar dinámicamente y adaptarse según las necesidades vuelven a este tipo de método en una herramienta muy útil a la hora de predecir el comportamiento del fuego.

Dependiendo del método, la búsqueda se puede centrar en un único conjunto de parámetros (solución única) o en varios conjuntos de parámetros (solución múltiple). El cómputo que exige estos métodos requiere de una gran capacidad de procesamiento, porque la búsqueda que se centra en las simulaciones puede llegar a superar las centenas de miles, por lo que tales métodos requieren de la ciencia computacional para su implementación.

Los métodos guiados por los datos, implementan cálculos de alto rendi-

miento que consisten principalmente en paralelizar los procesos de tal manera a reducir el tiempo y efectuar el mayor número de simulaciones posibles.

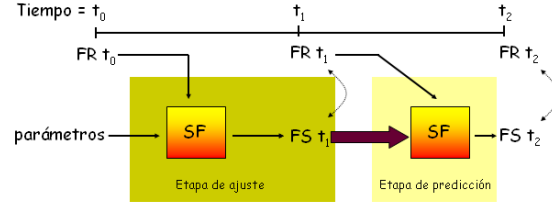


Figura 2.4: Predicción Guiada por los Datos

2.2.1. Solución Única

El método de solución única consiste en la búsqueda de un conjunto de datos que mejor represente el comportamiento del incendio. Para este caso las etapas de ajuste y predicción se definen de la siguiente forma [25]:

- Ajuste: En esta etapa es donde se aplicarán estrategias computacionales para realizar, a través del manejo efectivo de los datos y de la información que se dispone, una búsqueda efectiva que estará guiada por los mismos datos.

Durante la etapa de ajuste se evaluarán distintas combinaciones de valores para los parámetros, sintonizándolos hasta obtener una (o varias) combinaciones de valores para los parámetros, a fin de que se obtengan buenas simulaciones.

En la etapa de ajuste, se trata de determinar valores para los parámetros de entrada del simulador, tales que logren una buena simulación del instante de tiempo t_0 a t_1 . Para esto, se utiliza el simulador con la línea de fuego en el instante t_0 (FR_0) y los parámetros de entrada. El simulador retorna la línea de fuego simulada para el instante t_1 (FS_1). Esta línea de fuego es comparada con la línea de fuego real en el instante t_1 (FR_1) y se evalúa cómo ha sido la simulación utilizando ese conjunto de parámetros.

El resultado de las comparaciones, se utilizará como información para elegir mejores valores para cada uno de los parámetros de entrada y ésto se volverá a utilizar como entrada en el simulador para volver a evaluar el conjunto de parámetros “mejorado”.

Este proceso se repite hasta alcanzar un número de iteraciones preestablecido o hasta llegar a un umbral que determina un grado de bondad deseado en los resultados de las simulaciones.

- **Predicción:** Una vez concluida la etapa de ajuste, se elige el conjunto de parámetros que logre la mejor simulación para t_1 y se utilizará como entrada del simulador junto con la línea de fuego real en el instante t_1 (FR_1), para predecir el estado del incendio en el próximo instante de tiempo t_2 (FS_2).

Entonces, este método evalúa distintos conjuntos de valores y en base a los resultados de las simulaciones, se van mejorando dichos valores. El resultado de este método no está sujeto a una única combinación de parámetros sino a distintas combinaciones que se intentan mejorar paso a paso.

Es importante observar, que se utiliza un escenario que es bueno para un instante de tiempo, para predecir lo que pasará en el instante de tiempo posterior. Esto es debido a que los parámetros que describen bien un escenario hasta un instante de tiempo dado, se supone que también lo hará para el próximo instante de tiempo. En este punto, se considera que se agregará cierto grado de error en la predicción debido a la naturaleza cambiante de las características del ambiente.

Es importante considerar que es necesario utilizar algún criterio de selección, alguna forma de elegir o combinar los valores de los parámetros, que permita converger hacia combinaciones de valores que den buenas simulaciones. Para esto, se utiliza un algoritmo genético que intenta realizar de la mejor forma posible la búsqueda de una buena combinación de valores para los parámetros. En los próximos capítulos se dan las principales características de dicho algoritmo y cómo es utilizado en este entorno.

2.2.2. Solución Múltiple

En el universo de conjuntos de parámetros de entrada, existe la posibilidad de que no encontremos un solo conjunto de datos, que represente la mejor propagación del fuego en un instante dado. Por lo tanto, ¿Cómo seleccionar el mejor conjunto de parámetros para el paso siguiente?, actualmente los métodos de solución única no contemplan un mecanismo para seleccionar a un conjunto de parámetros que obtienen igual calidad de predicción, simplemente se ordenan de acuerdo al índice de error y se elige al primero de la lista.

- **Ajuste:** En esta etapa se realiza un barrido exhaustivo del espacio de búsqueda explorando un gran número de posibilidades con la combinación de valores de los parámetros. El método de solución múltiple, considera a todos los conjuntos de parámetros desde 0 hasta n . Entonces, todos estos conjuntos pueden representar distintas calidades de predicción que son solapadas para obtener un rango, valor, o índice

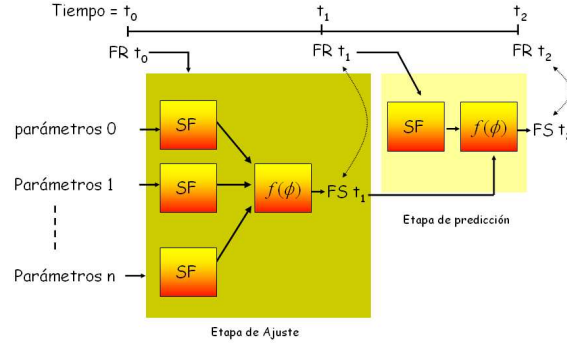


Figura 2.5: Solución Múltiple

para efectuar la predicción en el siguiente paso. En la figura 2.5 se observa cómo en la etapa de ajuste todos los conjuntos de parámetros son introducidos en el simulador de fuego SF , junto con el frente de fuego FR_{t_0} , cada resultado de la simulación es enviado al modulo $f(\phi)$ que es la encargado de integrar todas las simulaciones y construir el mapa simulado. Considerando que la calidad de cada simulación es medida como fitness (ϖ) y que la función $f(\varpi)$ presenta todos los resultados obtenidos entre 0 y un número n de simulaciones. El resultado final de la simulación consiste en encontrar un rango de integración de $[a, b]$ que refleje el comportamiento del frente de fuego en el instante FR_{t_1} , como muestra la ecuación 2.1.

$$f(\phi) = \int_a^b f(\varpi) d\varpi \quad (2.1)$$

- **Predicción:** Una vez concluida la etapa de predicción se realiza la simulación para el siguiente paso de tiempo, de t_1 a t_2 , aplicando el rango $[a, b]$ de la etapa de ajuste. Esta etapa se diferencia de la anterior principalmente porque la predicción exige adelantarse al comportamiento de fuego real en el tiempo FR_{t_2} . Por lo tanto, el resultado de la simulación se basa en un comportamiento anterior del incendio y la función $f(\phi)$ construye el mapa simulado a partir del rango de la etapa de ajuste. Por ejemplo, para un caso de 10 simulaciones, los resultados integrados de las simulaciones desde 0 a 5 son los que mejor semejanza tienen con el fuego real en la etapa de ajuste y entonces, para elaborar la predicción se vuelve a efectuar todas las simulaciones, pero esta vez para el instante de tiempo t_1 a t_2 y se integran todas las simulaciones del rango entre 0 a 5. De esta forma se construye el mapa simulado

sin la necesidad de compararlo con el FR_{t_2} . Pero una observación muy importante, es que esto no significa que una vez que el fuego llegue al tiempo t_2 no se pueda efectuar una comparación para ver que tan buena fue la predicción.

Para implementar ambos métodos (Solución Única y Solución Múltiple) se debe realizar un gran número de simulaciones, en donde se obtengan los conjuntos de parámetros que representen mejor o peor la propagación de incendio, pero no se debe confundir que por interactuar o procesar varios conjuntos de parámetros se clasifique como solución múltiple, porque la principal diferencia es que el de solución única descarta todas las soluciones que considera innecesarias para la predicción, quedándose con un solo conjunto de parámetros que utiliza para la etapa de predicción, mientras que el de solución múltiple tiene en cuenta todas las simulaciones efectuadas para realizar el cálculo de un valor o índice para la siguiente etapa en la predicción.

Capítulo 3

Métodos Guiados por los Datos

Estos métodos basan sus decisiones de acuerdo a como se van comportando los datos y esto es muy útil en entornos tan dinámicos como el de los incendios forestales. En este capítulo se describen dos métodos que son guiados por los datos. Primeramente se habla del método basado en un algoritmo evolutivo que se clasifica dentro de los métodos considerados de solución única descrito anteriormente, luego se describe el segundo método basado en un cálculo probabilístico dentro de los llamados métodos de solución múltiple. Ver figura 3.1.

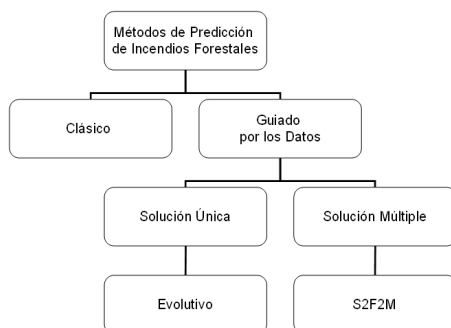


Figura 3.1: Taxonomía de los Métodos de Predicción de Incendios Forestales

Los métodos guiados por los datos se fijan principalmente en cual es la situación del incendio en un momento dado y confina que éste se comportará de una manera estable durante un cierto tiempo. Pero un cambio brusco en la situación del fuego necesitaría de nuevo una calibración en la etapa de ajuste y llevaría como mínimo un paso de tiempo para adaptarse a este nuevo comportamiento. Por ejemplo, se tiene un viento con una velocidad de 5 mph con dirección sur desde el tiempo, t_0 hasta el tiempo t_1 , el método puede adaptarse a estas condiciones y suponiendo que la velocidad y dirección del viento se mantengan constantes hasta el tiempo t_2 , el método brinda una

predicción fiable. Pero, si las condiciones varían bruscamente y el viento pasa de 5 a 10 mph y la dirección pasa de sur a norte, del tiempo t_1 a t_2 , cualquier ajuste en la etapa anterior será inútil, por lo que el método debe volver a adaptarse a estas nuevas condiciones para dar una predicción fiable en el tiempo t_3 , siempre y cuando las condiciones se mantengan.

A continuación, se describe cómo funcionan los métodos guiados por los datos que utilizan dos estrategias distintas para elaborar la predicción.

3.1. Método Evolutivo basado en Algoritmos Genéticos

El Método Evolutivo inventado a mediados de los 70 por John Holland, está inspirado en el modelo de evolución biológica y utiliza el principio de selección natural para resolver problemas complejos. Una definición bastante completa es la propuesta por John Koza:

“Es un algoritmo matemático altamente paralelo que transforma un conjunto de objetos matemáticos individuales con respecto al tiempo usando operaciones modeladas de acuerdo al principio de Darwin de reproducción y supervivencia del más apto, y tras haberse presentado de forma natural una serie de operaciones genéticas de entre las que destaca la recombinación sexual. Cada uno de estos objetos matemáticos suele ser una cadena de caracteres (letras o números) de longitud fija que se ajusta al modelo de las cadenas de cromosomas, y se les asocia con una cierta función matemática que refleja su aptitud” [26].

Primeramente, se debe conocer cuáles son las características principales de los algoritmos evolutivos en general. Este tipo de algoritmo simula la evolución genética natural. Se basa en evolucionar una población de individuos de forma iterativa y en dicha evolución sobreviven los individuos mejor adaptados al ambiente. Los individuos que sobreviven son los que tienen las mejores características en la población y las mismas se utilizan para generar nuevos individuos que formarán la próxima población. Por lo tanto, los nuevos individuos heredan las buenas características de los padres. Llevar esto a un entorno de programación implica definir qué es un individuo (al menos cómo está formado), definir una población de individuos y definir una función que determine cómo está adaptado un individuo a su ambiente. [25].

3.1.1. Descripción del Método

El método evolutivo consiste en incluir un mecanismo de búsqueda de parámetros óptimos para un incendio en un instante de tiempo $t_{(p-1)}$, en donde p es el tiempo de predicción. Entonces, los parámetros obtenidos en la búsqueda corresponden siempre a un paso de tiempo anterior para el cual se efectúa la predicción.

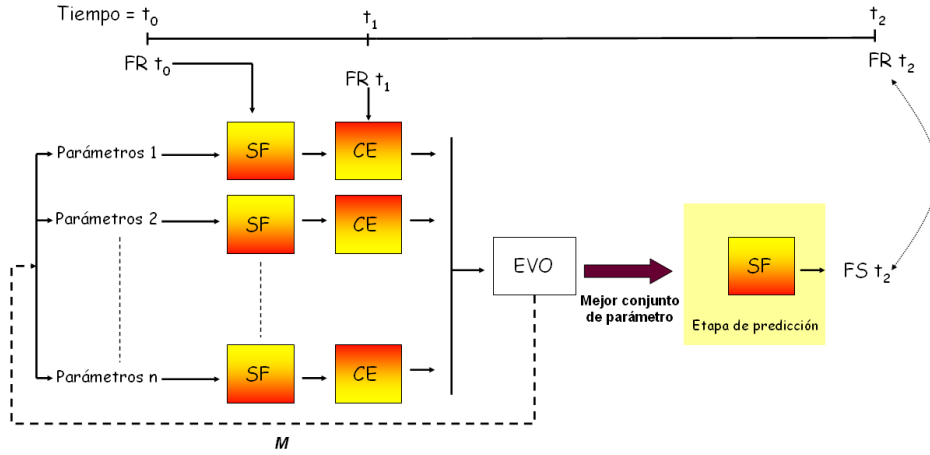


Figura 3.2: Metodo Evolutivo de Predicción de Incendios Forestales

A partir de la figura 3.2, se define el t_0 como el instante de tiempo de inicio de todo el método, entonces FR_{t_0} es el frente del fuego real en el momento que empezamos a observar la propagación del incendio y FR_{t_1} es el frente de fuego real en el instante t_1 , todos estos elementos corresponden a la etapa de ajuste. Se considera un espacio de búsqueda parcialmente desconocido, debido a que se conocen los rangos, con respecto a los valores de parámetros que pueden ser modificados por el método, pero existen constantes propias de la zona donde ocurre el incendio que deben ser obtenidas antes de iniciar la simulación. Los vectores de parámetros son evaluados mediante los resultados generados por el modulo SF (simulador de fuego) y comparados por el modulo CE (cálculo de error) que recibe los mapas generados del SF y los compara con el FR_{t_1} . Este proceso se efectúa M veces hasta que satisfaga la condición $error = 0$ ¹ o se cumpla el número de generaciones establecido por defecto en el módulo EVO , donde se encuentran todas las operaciones evolutivas.

Una vez que se realicen todas las operaciones sobre la población de conjuntos de parámetros con la cantidad M de iteraciones se selecciona el vector

¹La función de error se explica con mayor detalle en la siguiente sección.

de parámetros que tenga el valor de error más bajo y ésta es utilizada en la etapa de predicción dando el resultado FS_{t_2} (fuego simulado para el instante t_2). Entonces, cuando el incendio alcance el tiempo t_2 se puede comparar con el FR_{t_2} para medir cual fue la calidad de la predicción.

3.1.2. Implementación del Método

Para detallar la manera como trabaja el algoritmo evolutivo dentro del método se sigue la secuencia lógica establecida por la figura 3.3, propuesta John R. Koza en su libro Programación Genética [26], que trata sobre el esquema básico con que funciona un algoritmo evolutivo.

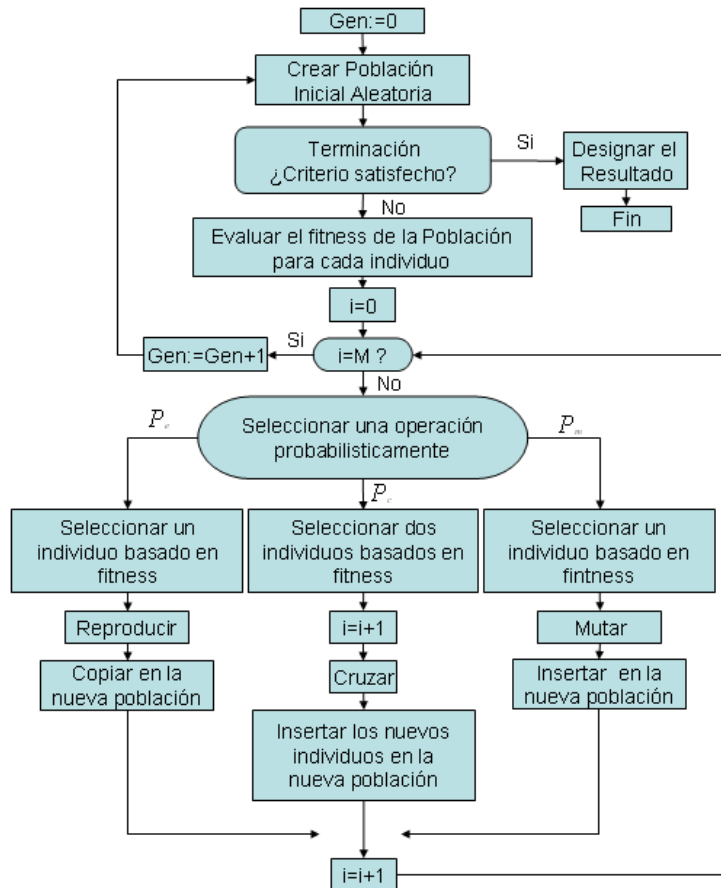


Figura 3.3: Diagrama de flujo convencional de un algoritmo genético

Primeramente, se debe definir según el esquema propuesto en la figura 3.4, el individuo y sus características con los cuales trabaja el algoritmo para luego crear una población aleatoria. Entonces, se toma como características del individuo los parámetros dinámicos, con que funciona el fireLib que son: M1 (representa la carga del combustible), M10 (humedad del combustible muerto en 10 horas), M100 (humedad del combustible muerto en 100 horas), MHERB (humedad del combustible herbáceo vivo), SAVR (razón entre superficie y volumen), DEPTH (profundidad del lecho del combustible), MEXT (contenido de la humedad de extinción del combustible), WINDSPD (velocidad del viento), WINDDIR (dirección del viento). Una observación importante, es que el método no efectúa ninguna operación sobre algunas variables estáticas que se mantienen constantes durante todo el proceso, por lo que no forman parte de las características del individuo a los que se le aplicarán las operaciones del algoritmo evolutivo que son: el SLOPE (pendiente del terreno), el ASPECT (orientación de la pendiente según al ángulo con el que se observa) y el MODELO (el fireLib trabaja con doce clases de modelos que se clasifican según el tipo de vegetación que compone el terreno donde se efectúa el incendio).

Carga	M1	M10	M100	MHERB	SAVR	DEPTH	MEXT	WINDSPD	WINDDIR
0.375924	0.146256	0.812345	0.225166	1.147839	873.7777	4.403738	0.442637	6.813624	138.889

Figura 3.4: Individuo de la población

A partir del individuo que representa un conjunto de parámetros se construye una población aleatoria, figura 3.5, suponiendo que no se conocen los valores de dichos parámetros, pero sí el rango en que el valor puede variar. Por ejemplo, el viento puede tener una dirección entre 0 y 360 grados, la velocidad entre 0 a 15 mph (a una velocidad mayor de viento pasa de ser moderado a turbulento en donde el modelo de ROTHERMEL [10] considera imprevisible el comportamiento del incendio), la M1 entre 0 y 0.4 (a mayor valor la humedad no permite la propagación del incendio). Para generar cada valor de un parámetro se define la ecuación 3.1, en donde $f(r)$ es una función randómica basada en una distribución uniforme para variable aleatoria continua, que devuelve valores entre $0 \leq x < 1$ multiplicándolo por un rango válido que se define como φ para el límite inferior y ϖ para el límite superior de cada parámetro x_i .

$$x_i = f(r) * (\varpi - \varphi) + \varphi \quad (3.1)$$

Una vez generada la población aleatoria se define la función por el cual se evalúa la calidad de cada individuo, se dice que un individuo está mejor

0.375924	0.146256	0.812345	0.225166	1.485.079	173.893.860	3.800.560	0.829998	6.813.624	135.000
1.348.169	0.318893	0.267325	0.406499	2.359.029	2.824.924.316	1.338.930	0.244621	7.279.422	130.000
1.348.169	0.318893	0.267325	0.406499	1.444.140	2.824.924.316	1.338.930	0.244621	7.279.422	132.000
1.790.955	0.275587	0.508409	0.414521	1.981.330	1.027.904.785	5.098.024	0.754927	7.176.535	136.000
0.115794	0.350956	0.098592	0.615440	1.889.236	2.816.164.795	2.036.088	0.420297	7.850.167	135.000
1.907.543	0.302298	0.444973	0.002878	1.133.699	3.173.971.924	2.632.373	0.618788	7.030.702	131.000
0.358069	0.163245	0.945372	0.194335	0.604616	1.835.489.746	2.529.896	0.223979	7.337.354	133.000
0.196998	0.112388	0.912107	0.341670	0.435321	1.886.948.486	5.840.542	0.408081	6.591.928	137.000
1.327.518	0.164227	0.990717	0.424103	3.844.979	1.284.650.024	1.356.648	0.637660	7.125.043	129.000
1.358.298	0.237787	0.254425	0.247351	3.689.048	2.824.924.316	1.338.930	0.244621	6.777.987	139.000

Figura 3.5: Ej. Población de 10 individuos

adaptado al ambiente cuando el mapa simulado es similar al mapa del fuego real (ver figura 3.6) y está definida por la función de error representada por la ecuación 3.2.

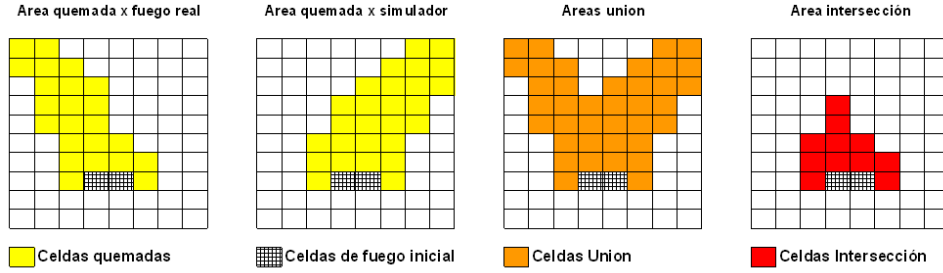


Figura 3.6: Cálculo de Error

$$Error = \frac{(Celdas \cup -CeldasInicio) - (Celdas \cap -CeldasInicio)}{CeldasReal - CeldasInicio} \quad (3.2)$$

Esta función permite saber las celdas equivocadas por exceso o falta de quemado en la simulación y si la función $Error = f(x)$, se puede definir la función objetivo del algoritmo evolutivo como la de minimizar la función $f(x)$. Esto quiere decir, a un menor valor de error, mejor adaptado está el individuo al ambiente, por lo tanto, mayor calidad tendrá la simulación con dicho individuo en ese instante. Puede presentarse la situación en que cuando se crea la población aleatoria y aunque las posibilidades son bastante bajas, uno de los individuos creados tenga un $error = 0$, consecuentemente este individuo es seleccionado para la etapa de predicción sin continuar con el proceso cumpliéndose el criterio de satisfacción.

Una vez creada la población aleatoria se pasa a las operaciones propias del algoritmo evolutivo (Selección, Cruzamiento y Mutación), primero se inicializa la variable contador de iteración i y se compara con la variable M que es el número de generaciones que se define para evolucionar la población. El

valor de M se ha establecido en 5 debido a que a una mayor iteración el valor de error se mantendrá estable, tal como se muestra en el gráfico 3.7, y como iterar más significa más cómputo con un costo de tiempo, se recomienda el valor de M entre 5 a 10.

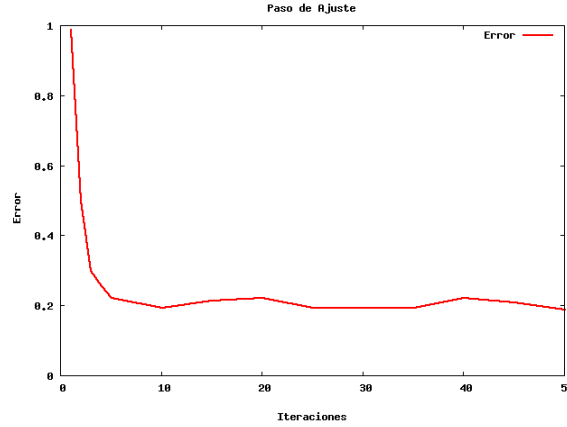


Figura 3.7: Comportamiento del error según el número de iteraciones

Los siguientes procesos de algoritmo siguen un esquema en base a la probabilidad, que representan las operaciones básicas del principio de evolución natural:

- Elitismo (P_r): Esta operación se rige por un valor de probabilidad fijada por defecto, con un valor del 2% de la población total que son ordenados según su puntuación, de esta forma se copian los n primeros a la siguiente generación de población para garantizar que los individuos que tengan mayor calidad se mantengan como parte de la población.
- Cruzamiento (P_c): La operación para seleccionar los individuos que son cruzados utiliza una técnica denominada *Rueda de Ruleta* [25] en donde cada individuo de la población, según su puntuación tiene una mayor proporción de la ruleta lo que permite a los individuos con mayor calidad tener más probabilidades de ser elegidos como padres. El tipo de cruce efectuado es el *segmentado*, tal como muestra la figura 3.8, en donde se elige un punto de cruce que divide al individuo en dos grupos de genes y éstos son intercambiados para generar dos nuevos individuos.
- Mutación (P_m): Se utiliza la mutación por intercambio en donde un gen del mismo tipo es seleccionado de uno de los padres y puesto al

nuevo individuo, figura 3.8, el valor de probabilidad configurado es muy bajo, igual a 1 %, esto se debe a que la mutación en la naturaleza es un fenómeno poco frecuente pero necesario para que los individuos se adapten a nuevas condiciones, un valor de mutación muy alto puede caer en máximos y mínimos locales lo que reduce la variabilidad de la población restringiendo la exploración de parámetros a un espacio limitado de búsqueda.

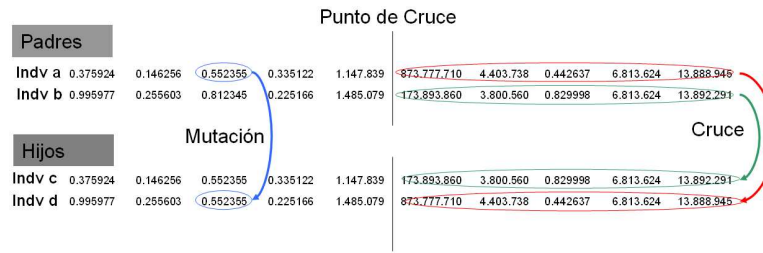


Figura 3.8: Cruce y Mutación

Luego de realizar todas las operaciones, y la nueva población es formada, todo el proceso es repetido hasta alcanzar el número de generaciones deseado.

3.1.3. Ventajas

- En un universo tan complejo donde se utilizan simuladores típicamente paramétricos y las variables de entrada tienen un alto grado de dificultad para medirlas y obtenerlas de manera rápida y sencilla, este método obtiene importantes ventajas con respecto a las metodologías tradicionales. Como se planteó anteriormente, una mala entrada de parámetros produce una mala salida y las métricas utilizadas por la física necesitan un alto nivel de precisión. Este algoritmo es potencialmente una de las mejores técnicas de optimización de parámetros que puede dar respuestas a estas dificultades a un nivel aceptable de confianza.
- Y como el tiempo es un factor gravitante en la metodología, el algoritmo evolutivo es esencialmente paralelo permitiendo realizar varios procesos simultáneos y con una fácil adaptación a las modernas arquitecturas masivamente paralelas, disminuyendo considerablemente el tiempo de respuesta para una predicción. Además, existen diversas técnicas de programación e implementación para convertir el método en una aplicación de tiempo real.

- El método no necesita de una estructura compleja para ser implementado, más bien consiste en componer estructuras simples que interactúen entre sí dejando que la evolución solucione el problema. Asimismo, es un método capaz de encontrar soluciones aún cuando las condiciones iniciales del incendio cambien con el tiempo.
- El método es independiente del simulador subyacente, esto produce una mayor flexibilidad a la hora de implementarlo en cualquier región, debido a que existen simuladores de fuego especialmente diseñados para zonas geográficas específicas.
- La paralelización de esta metodología encaja perfectamente dentro del paradigma Master/Worker.
- En principio, el método evolutivo puede ser mejorado incorporando conocimiento al algoritmo, como es la propuesta presentada por Mónica Denham en su memoria de trabajo “PREDICCIÓN DE INCENDIOS FORESTALES BASADA EN ALGORITMOS EVOLUTIVOS GUIADOS POR LOS DATOS” [25]. De esta manera se puede converger en un resultado más preciso y rápido.

3.1.4. Desventajas

- El método trata de explorar las regiones más prometedoras de un enorme espacio de posibilidades. Al encontrar una zona con pesos altos, ésta se explora más a fondo. Pero hay que evitar que el algoritmo se estanque en una determinada zona, produciendo multitud de vectores de parámetros muy parecidos.
- Pueden tardar mucho en converger, o no converger en absoluto, dependiendo en cierta medida de los parámetros que se utilicen: tamaño de la población, número de generaciones, probabilidad de mutación o cruzamiento.
- Si bien el método utiliza el mejor resultado obtenido esto no garantiza que sea la solución óptima y depende en gran medida de la zona de espacio de búsqueda explorado.
- El método puede encontrar el mejor conjunto de parámetros para la etapa de ajuste, pero en algunos casos este mismo set de parámetros no obtiene una buena calidad en la predicción, esto podría ser mejorado si se tiene en cuenta los demás individuos que en este método son desechados y basan su predicción en un solo individuo.
- A un mayor número de individuos de la población mayor tiempo de simulación, esto se debe a que el número de individuos de la población

se mantiene constante durante todas las iteraciones, es recomendable alguna técnica que deseche a individuos que sean inservibles, por ejemplo, algunos individuos pueden superar ciertos umbrales de valores que imposibiliten la propagación del fuego.

3.2. Método Estadístico - S^2F^2M (Statistical System for Forest Fire Management)

El término estadística tiene su raíz en la palabra Estado. Surge cuando se hace necesario para sus intereses cuantificar conceptos. En la mayoría de los casos, esta cuantificación se hace en función de unos fines económicos o militares. El estado quiere conocer censo de personas, de infraestructura, de recursos en general, para poder obtener conclusiones de esta información.

Actualmente la estadística es una ciencia. No es ya una cuestión reservada al estado. Se puede decir que se encuentra en la totalidad del resto de las ciencias. La razón es clara: por una parte, la estadística proporciona técnicas precisas para obtener información, (recogida y descripción de datos) y por otra parte, proporciona métodos para el análisis de esta información.

Formalmente la estadística está definida como un conjunto de métodos para planear experimentos, obtener datos, organizar resúmenes, analizarlos, interpretarlos y de ellos extraer conclusiones. Pero este método, va un poco más allá que la estadística denominada descriptiva que se centra en los métodos de recolección, descripción, visualización y resumen de datos originados a partir de los métodos de estudio. Sino que también utiliza la inferencia estadística, que se dedica a la generación de los modelos, inferencias y predicciones asociadas a los fenómenos en cuestión, teniendo en cuenta lo aleatorio e incertidumbre en las observaciones.

En cualquier metodología heurística, la estadística es una herramienta importante para demostrar la hipótesis planteada y obtener conclusiones a partir de los datos, los incendios forestales no son la excepción, pero hay que tener bien claro que el método no basa su predicción en históricos de datos, aunque el modelo permite introducir conocimiento, entonces utilizando un simulador de fuego y generando una cantidad n de conjuntos de parámetros que representan los escenarios posibles en los cuales el fuego puede propagarse, se construye un mapa de probabilidades del comportamiento del incendio como se explica detalladamente más adelante.

¿Porqué metodología estadística de predicción de incendios forestales?

Porque el método utiliza el cálculo de probabilidad para elaborar la predicción y la cuestión planteada, en un universo determinista basado en conceptos muy parecidos a la mecánica Newton, donde no existe probabilidad si se conocen todas las condiciones. Un ejemplo de este concepto sería el de una ruleta, si la fuerza de la mano y el periodo de esta fuerza es conocido, entonces el número donde parará la bola será seguro de predecir. Natural-

mente, esto también supone el conocimiento de la inercia y la fricción de la ruleta, el peso, lisura y redondez de la bola, las variaciones en la velocidad de la mano durante el movimiento y así sucesivamente. De igual manera, en la física de propagación de incendios forestales, la teoría probabilística puede resultar mucho más práctica que aplicar los conceptos de la mecánica newtoniana que necesitan de gran cantidad de variables y mediciones demasiado complejas.

3.2.1. Descripción del Método

En teoría existen dos maneras de obtener datos sobre un evento. Una manera es por medio de un estudio observacional, en tal caso se debería ver como se comporta un incendio real y medir todas las variables en el cual se desarrolla. La otra manera es por medio de experimentos diseñados en donde nosotros podemos intervenir sobre los valores de dichas variables y ver cuál es el comportamiento del incendio (para ello, es indispensable un laboratorio con los equipos necesarios para reproducir las condiciones ambientales y topográficas). De las dos formas, la primera es muy difícil de efectuar por lo complejo de medir dichas variables, además rara vez un incendio presenta el mismo comportamiento dos veces debido a que las variables cambian. La segunda, mucho más factible y gracias al avance en la ciencia computacional, utilizando un simulador de incendio forestal, como el fireLib, se puede efectuar varios tipos de incendios e interactuar con la variables para observar el comportamiento del fuego, de esta manera el laboratorio es la propia computadora. Consecuentemente, esto permite elaborar estrategias basadas en probabilidades para efectuar predicciones por medio de experimentos factoriales [27].

Este método utiliza la misma estrategia que el método evolutivo, pero al contrario de utilizar un algoritmo genético se vale del factor de probabilidad (k) para elaborar los mapas de predicción. Por lo tanto, primero se observa y luego se predice, tal como se explicó en la sección anterior de métodos guiados por los datos. Por consiguiente, a partir de la figura 3.9 se define el tiempo t_0 como el inicio de todo el método y entonces el FR_{t_0} , FR_{t_1} y FR_{t_2} son el frente de fuego real en los instantes t_0 , t_1 y t_2 . El método se encuentra dividido de la siguiente manera:

Etapas de Ajuste: La etapa de ajuste se divide en dos grandes bloques y de acuerdo a los datos efectúa el cálculo del factor a ser utilizado en la etapa de predicción.

- Simulación Combinatoria (SC): Este bloque es el que utiliza al simulador de fuego (SF) como herramienta para elaborar el mapa de probabilidades para el instante t_0 a t_1 . En este punto se realiza un barrido exhaustivo del espacio de búsqueda de parámetros y cada resultado del SF es integrado en el módulo de Simulación Estadística (SE) para

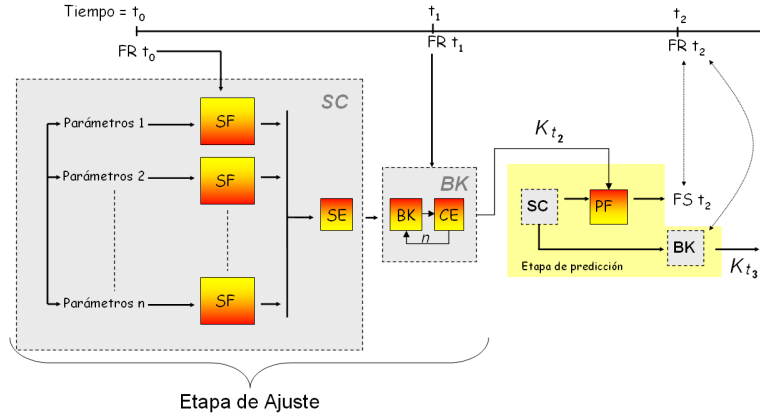


Figura 3.9: Método Estadístico de Predicción de Incendios Forestales

elaborar el mapa de probabilidades sobre el terreno y ser enviado a la siguiente etapa.

- Búsqueda de K_{ign} (BK): El factor K_{ign} por su siglas en inglés (Key Ignition Index) está definido como el valor que reproduce la probabilidad de ignición del terreno que mejor se ajusta a la propagación del incendio real y recibe como entrada el mapa de probabilidades del bloque SC. Obsérvese que BK necesita del mapa de propagación del fuego real como entrada y utiliza el módulo de Cálculo de Error (CE); esta función es la misma utilizada para el método evolutivo y representado por la ecuación 3.2, para comparar el mapa de probabilidades con el fuego real y medir qué valor de K_{ign} presenta menor error en comparación con el FR_{t_1} , esto se efectúa n veces (donde n es el número de escenarios simulados por SC) hasta encontrar el valor de $error = 0$ o bien el valor de error más bajo.

Etapa de Predicción: En esta etapa se incluyen los dos bloques SC y BK utilizados en la etapa de ajuste con un nuevo módulo que es el encargado de efectuar la Predicción del Fuego (PF). Pero, se debe tener en cuenta que el bloque SC, a diferencia de la etapa de ajuste, genera el mapa de probabilidades para el instante t_1 a t_2 y que recibe como entrada el módulo PF, en donde según el factor K_{ign} de probabilidades recibido de la etapa de ajuste se genera el mapa simulado para el instante t_2 . Una observación importante es que el bloque BK efectúa el cálculo de la K_{ign} para el instante t_3 , de esta manera el mapa de probabilidades generado para la predicción de t_1 a t_2 es aprovechado para efectuar el proceso de ajuste para el siguiente paso de tiempo t_3 .

3.2.2. Implementación del Método

Al igual que el método evolutivo se sigue una estructura lógica para su implementación como se encuentra representada por la figura 3.10 [27].

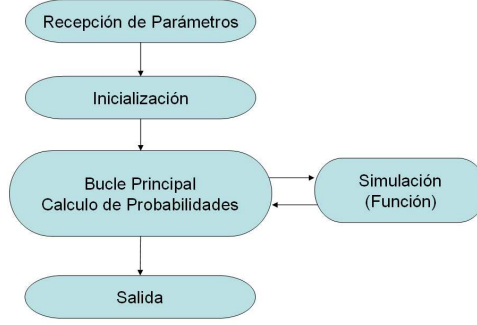


Figura 3.10: Ej. Estructura de S^2F^2M

Para el funcionamiento del método es necesario definir los rangos para cada tipo de parámetro más un valor de incremento, esto es porque el método debe crear diversos grupos de parámetros con valores distintos, cada conjunto de parámetros representa un escenario, tal como muestra la figura 3.11. Estos valores especifican el espacio de búsqueda a ser explorado y la variedad de situaciones posibles dentro de un incendio real. Esta primera etapa corresponde a la *Recepción de Parámetros* dentro de la figura 3.10 y es la que marca la complejidad de todo el proceso de búsqueda exhaustiva.

Para cada parámetro denominado $Parametro_i$ se define un umbral mínimo (U_{min}) y máximo (U_{max}) que representa el intervalo en el cual se moverá el valor con ayuda de un factor de incremento, llamado $incremento_i$. Para cada parámetro es posible obtener un número C_i que representa el número de valores que puede tomar dicha variable definido como dominio de cardinalidad. Entonces para cada $Parametro_i$ se tiene la ecuación 3.3.

$$C_i = \frac{(U_{max} - U_{min}) + incremento_i}{incremento_i} \quad (3.3)$$

Una vez que se obtiene la cardinalidad de cada parámetro es posible calcular la cantidad de escenarios posibles de acuerdo a un número p de parámetros como muestra la ecuación 3.4:

$$n^{\circ}escenarios = \prod_{i=1}^p C_i \quad (3.4)$$

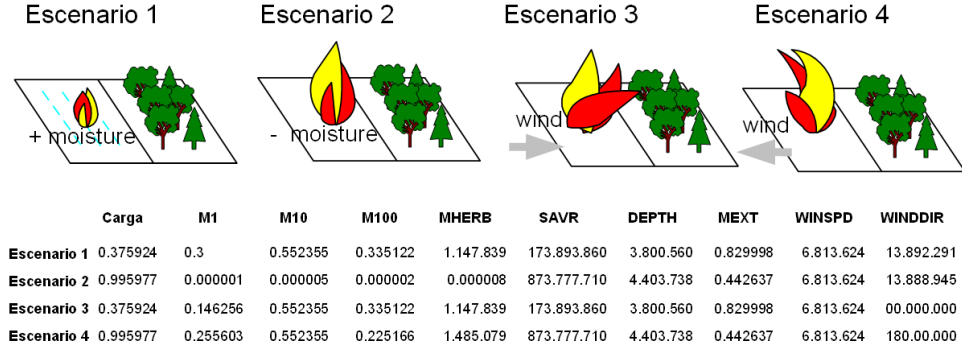


Figura 3.11: Ej. Escenarios de simulación

En la figura 3.12, se observa el ejemplo de un archivo de parámetros que contiene en la primera columna el nombre del parámetro, en la segunda se puede apreciar el valor del U_{min} en la tercera columna el U_{max} y en la cuarta columna el factor de $incremento_i$.

#Model	7	7	1
#WindSpd	0.67	12.63	.67
#WindDir	0.0	360.0	45.0
#Slope	18	18	1.0
#Aspect	180.0	180.0	1.0
#M1	0.1	0.4	0.1
#M10	0.1	0.3	0.2
#M100	0.1	0.5	0.2
#Mherb	0.1	0.4	0.2

Figura 3.12: Archivos de parámetros de S^2F^2M

Potencialmente todos los parámetros pueden ser variados para realizar una búsqueda más exhaustiva, pero es recomendable mantener la pendiente (SLOPE) y la orientación (ASPECT) del terreno con un valor constante, de manera a simplificar y dirigir la búsqueda sobre aquellos parámetros cuyos valores son difíciles de obtener debido a su comportamiento dinámico o mala lectura.

Luego de la recepción de los parámetros, corresponde la etapa de inicialización de todas las variables, se calcula el número de escenarios posibles, la creación de las matrices que contienen el mapa de fuego inicial y la matriz de probabilidades donde se guardarán los contadores de ignición para cada porción del terreno. En esta etapa, se procesa el contenido leído del fichero de entrada, más el mapa de fuego inicial definida en la figura 3.9, como FR_{t0} .

La carga más pesada en cuanto a proceso de todo el método se encuentra en el bucle principal, que es el encargado de elaborar el mapa de probabilidades sobre la matriz del terreno. En esta sección cada celda de la matriz corresponde a una porción del terreno y lo que se quiere saber es si dicha celda fue quemada o no para un escenario y un intervalo de tiempo específico. Entonces, si tenemos que n es el número total de escenarios y n_A es el número de escenarios en donde la celda fue alcanzada por el fuego, se puede calcular la probabilidad de ignición para esa porción del terreno, como muestra la ecuación 3.5:

$$P_{ign}(A) = \frac{n_A}{n} \quad (3.5)$$

Consecuentemente, cada conjunto de parámetros es enviado a la función de simulación (ver figura 3.10) y esta devuelve una matriz con las celdas alcanzadas por el fuego, que es acumulada en la matriz de probabilidades como muestra la figura 3.13. No necesariamente las celdas que contienen el

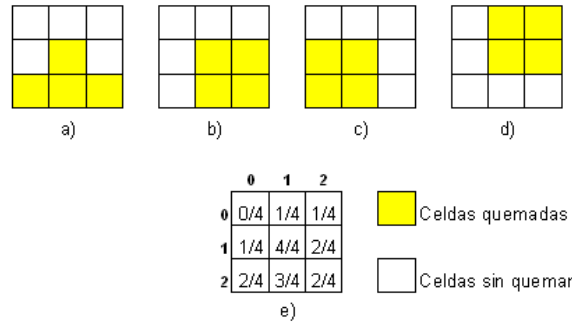


Figura 3.13: Ejemplo de Matriz de Probabilidades

mismo valor de probabilidad P_{ign} fueron originadas por el mismo conjunto de parámetros. Es muy difícil de comprobar qué conjuntos de parámetros exactamente ha generado un valor de probabilidad dentro del mapa.

En las figuras 3.13.a, 3.13.b, 3.13.c y 3.13.d, se pueden apreciar las celdas que fueron quemadas por conjuntos de parámetros distintos y cada una de ellas representa un escenario diferente. Entonces, cada escenario es evaluado

e integrado en una matriz con la $P_{ign}()$ por cada celda. Por ejemplo, se observa que la celda (0,0), no ha sido quemada en ninguno de los 4 escenarios por lo que tiene una probabilidad igual a 0/4, las celdas (0,1) y (0,2) han sido quemadas en un solo escenario correspondiente al d y lo mismo ocurre para la celda (1,0) que corresponde al escenario c , todos éstos con una probabilidad de 1/4. La $P_{ign}()$ de la celda (1,1) se obtiene a partir de solapar las 4 matrices pues han sido quemadas en todos los casos con una probabilidad de 4/4 y la celda (2,1) con una probabilidad del 3/4 ya que ha sido quemada por los escenarios a , b y c . Cuando el número de escenarios es demasiado elevado es una tarea prácticamente imposible determinar que escenario ha generado valor de P_{ign} para cada celda.

Por lo tanto, una vez que se tiene el mapa de probabilidades se llega al punto fundamental de todo el método, que consiste en encontrar el factor K_{ing} que reproduce el valor de probabilidad que mejor se adapta al comportamiento del fuego real. De este modo, se compara el mapa de probabilidades con el mapa del fuego real en busca del K_{ign} , que satisfaga la condición de la ecuación 3.6

$$\{x : P_{ign}(x) \geq \frac{K_{ign}}{n} | K_{ign} \in N\} \quad (3.6)$$

en donde n es igual al número de escenarios y P_{ign} es la función de variación desde K_{ign}/n hasta 1, esto representa, por ejemplo, que x es el conjunto de celdas que se han quemado por lo menos K_{ign} veces.

Para graficar mejor este concepto se plantea la lógica de conjuntos, sea entonces cuatro conjuntos de celdas a partir del ejemplo de la figura 3.14.a con las siguientes definiciones:

$$\begin{aligned} A &= \{x/x \text{ tiene } K_{ign} \geq a\} \\ B &= \{x/x \text{ tiene } K_{ign} \geq b\} \\ C &= \{x/x \text{ tiene } K_{ign} \geq c\} \\ D &= \{x/x \text{ tiene } K_{ign} \geq d\} \end{aligned}$$

Siendo a , b , c , y d , cualquier número entero positivo que reproduce la probabilidad de cada conjunto de celdas y por otro lado, tenemos el mapa de fuego real para el instante t_1 (FR_{t_1}), a partir de aquí se define el conjunto R para las celdas quemadas en el fuego real:

$$R = \{x/x \text{ es una celda quemada por el fuego real}\}$$

Todos estos conjuntos presentan la siguiente característica (teniendo en cuenta que la probabilidad dentro del método tiene un comportamiento creciente $a > b > c > d$). Entonces,

$$A \subseteq B \subseteq C \subseteq D$$

donde el conjunto D es el que mayor celdas contiene e incluye a todos los demás conjuntos y el conjunto A el que menor celdas contiene y esta incluido en todos los demás conjuntos, si se observa el gráfico de la figura 3.14.a, esta propiedad se ve reflejada de la siguiente manera en la figura 3.14.b, a un mayor valor de K_{ign} menor superficie abarca y a un menor valor K_{ign} mayor valor de superficie. Por lo tanto, el K_{ign} y S son inversamente proporcionales.

Si se aplica la operación de intersección \cap entre el conjunto R y cada uno de los conjuntos (A, B, C, D) , se puede definir la función ω_{max} como la de maximizar el número de elementos de la intersección entre un conjunto ya sea A, B, C, D y el conjunto R , de esta manera se obtiene:

$$|R \cap A| = x_a; |R \cap B| = x_b; |R \cap C| = x_c; |R \cap D| = x_d$$

Si:

$$x_b > x_a > x_c > x_d$$

Por lo tanto, la función $\omega_{max}(x) = x_b$ por lo que se concluye que el conjunto B es el que contiene mayor número de celdas en común con el conjunto R y el valor del K_{ign} para dicho conjunto es el que presenta el valor que reproduce la mejor probabilidad con el mapa real. Para el método, una probabilidad alta no significa que sea el mejor valor de K_{ign}

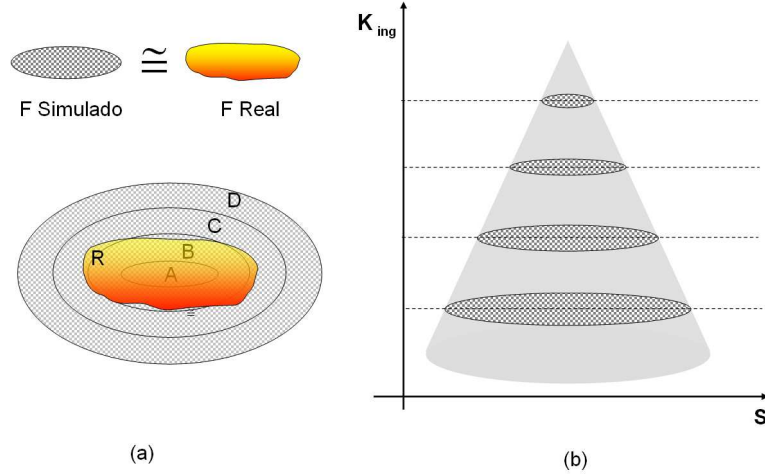


Figura 3.14: Esquema del método S^2F^2M

Como se plantea en la analogía del cono de probabilidades de la figura 3.14.b, el valor K_{ign} nos indica la altura del corte que se debe hacer sobre el mapa de probabilidades. Esta K_{ign} es la que será utilizada en la siguiente etapa para elaborar el nuevo mapa de predicción.

El siguiente paso es elaborar de nuevo la matriz de probabilidades, pero

esta vez para el siguiente paso de tiempo, y en esta matriz sólo son seleccionadas las celdas (x) que cumplan la condición de la ecuación 3.6. De esta forma se construye el mapa de predicción de la propagación del incendio. Otra observación importante es que el método no indica en qué momento exactamente la celda será alcanzada por el fuego, pero sí el periodo de tiempo en que ésto ocurrirá.

3.2.3. Ventajas

- Una de las principales ventajas que presenta el método es la capacidad de cubrir una gran variedad de condiciones que se pueden presentar en una situación real, esto es denominado como búsqueda exhaustiva del espacio de valores que pudiera tomar los parámetros que representan dichas condiciones.
- Los resultados experimentales obtenidos en las investigaciones previas de este trabajo [27], demuestran que el método se adapta perfectamente a los cambios de condiciones, redireccionando la simulación hacia un resultado más semejante al comportamiento del fuego real, haciéndolo independiente de la necesidad de medir los cambios de las variables dinámicas, esto no significa que obtener dichas mediciones no sean de gran utilidad, pero es casi imposible obtenerlos en un tiempo razonable por lo que muchas veces es mejor descartarlos.
- El simulador es una herramienta independiente dentro del método que permite utilizar con una gran variedad de simuladores y esto permite potenciar los resultados obtenidos por un simulador convencional.
- Todos los simuladores de propagación de incendios forestales son dependientes de la calidad de datos con que se cuente en el momento de realizar la predicción, por lo tanto, una estrategia guiada por los datos con herramientas probabilísticas es de gran utilidad a la hora de suplir la necesidad de detalle que exige la física de los modelos implementados en los simuladores.
- El método encaja perfectamente dentro del paradigma master/worker, uno de los modelos clásicos del paralelismo por lo que su implementación es relativamente sencilla en una arquitectura paralela como un Cluster o Grid.

3.2.4. Desventajas

- El modelo exige gran capacidad de cómputo para reducir el tiempo de respuesta de una predicción, por lo que su rendimiento es muy dependiente de la capacidad del Cluster o Grid donde se implemente.

- La principal virtud del método, la búsqueda exhaustiva es también su principal problema a la hora de implementarla como una aplicación de tiempo real, debido a la cantidad de simulaciones necesarias, esto puede llegar a superar el millón de simulaciones y dependiendo del tipo de simulador cada simulación individualmente tiene una duración variable.
- Puede llegar a producirse un gran número de simulaciones en donde la combinación de parámetros provoca que el incendio no se propague siendo una carga innecesaria de cómputo.

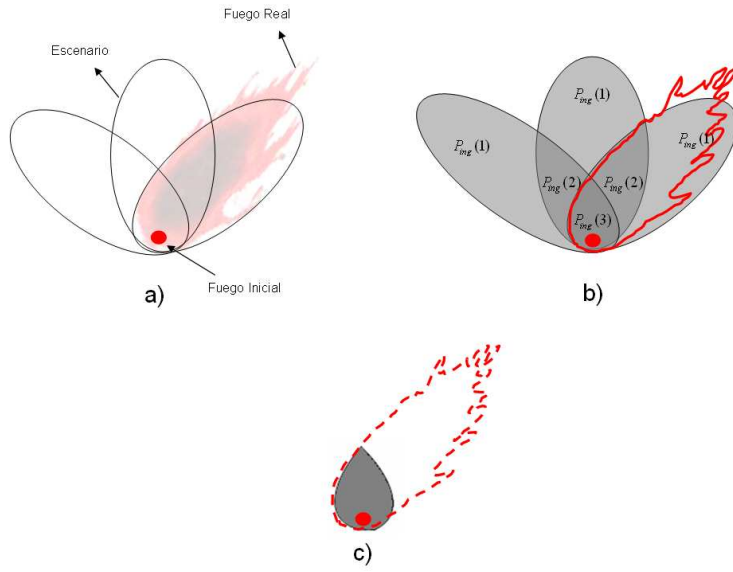


Figura 3.15: Resultado de baja calidad

- En algunos tipos de incendios el método puede llegar a producir un mapa muy diferente al caso real, pese a que uno de los escenarios obtuviera un mapa muy semejante como muestra la figura 3.15.a. Esto se debe a que al solapar los escenarios que se obtuvieron como resultado del barrido exhaustivo del espacio de búsqueda, producen un mapa de probabilidades como la figura 3.15.b en donde la probabilidad que mejor se adapta es la $P_{ign}(3)$ que da como resultado el mapa de la figura 3.15.c. Esto se traduce a un valor de error muy alto en la etapa de ajuste, pero más adaptable a los cambios de condiciones durante la siguiente etapa. Pero si la situación en donde se desarrolla el incendio se mantiene estable en el transcurso del tiempo bajo estas circunstancias S^2F^2M no podrá converger en una buena predicción.

Capítulo 4

*SAPIFE*³ - Sistema Adaptativo para la Predicción de Incendios Forestales basado en Estrategias Estadístico-Evolutivas

La complejidad del problema tratado en el área de incendios forestales, hace necesaria cada día más, una gran capacidad de cálculo. Es un hecho bien establecido que para algunos de estos problemas (métodos numéricos, simulación, optimización, etc.), la computación paralela ofrece soluciones eficientes [11], en donde S^2F^2M puede ser implementado de una manera factible, pero esto en la práctica requiere ser ejecutado en un cluster de gran tamaño. Sin embargo, las supercomputadoras paralelas continúan sin estar al alcance de muchas organizaciones debido a sus elevados costes en hardware, mantenimiento y programación. Otra solución puede ser ejecutar el sistema dentro de entornos geográficamente distribuidos como son actualmente los Grids Computing, esto presenta una alternativa muy interesante, pero la restricción de tiempo real en este entorno es una variable que puede representar algunos problemas, que llevarán un poco más de tiempo en ser salvados. Otra alternativa, sería aplicar alguna técnica de optimización de parámetros que reduzca el número de simulaciones necesarias de S^2F^2M , de tal forma a que la capacidad de cómputo no sea una restricción fuerte dentro de esta metodología, y en cuanto se refiere a técnicas de optimización, los algoritmos evolutivos son una referencia dentro este campo.

De estas alternativas surge la propuesta de una integración de métodos, con el objetivo de aprovechar las fortalezas de ambas metodologías, entonces la idea base a este planteamiento es la de optimizar las entradas de S^2F^2M con el algoritmo evolutivo *AG*, a fin de reducir el número de simulaciones

necesarias y que esta reducción no presente una disminución considerable de calidad en la predicción. Pero, por supuesto que antes se debe estudiar la complejidad de esta propuesta y su viabilidad.

En el capítulo anterior, se analizaron las ventajas y desventajas de ambos métodos, ahora se debe estudiar si esta alternativa presenta problemas irresolubles. Y es aquí donde surgen las preguntas ¿son compatibles ambos métodos? o en todo caso, ¿pueden ser complementarios?. Las respuestas a estas preguntas requieren analizar las bases teórico-prácticas de las metodologías, para que la propuesta sea coherente con los principios básicos de ambos.

4.1. Bases Teóricas y Prácticas de la Propuesta

Los métodos guiados por los datos tienen como base principios comunes “*Conocimiento y Adaptabilidad*”, estos principios son muy simples conceptualmente. Dentro de estas metodologías, el conocimiento implica dos cosas una, es medir las condiciones en donde se desarrolla el incendio en tiempo real (velocidad del viento, dirección del viento, humedad del combustible, pendiente del terreno, etc.) y otra, es saber cuál es el comportamiento del incendio (esto se mide por la forma geométrica que el fuego va adquiriendo). La adaptabilidad consiste en modificar el comportamiento del simulador a una nueva situación, basándose en información que advierte posibles cambios que pueden ocurrir en las condiciones ambientales, una cualidad demasiado útil y necesaria en el entorno tan dinámico en el que se desarrolla un incendio forestal.

Lo interesante de los dos tipos de conocimiento, ya sea de las condiciones o comportamiento, es que a partir de una se puede determinar la otra. Esto quiere decir, que conociendo las condiciones ambientales se puede simular el comportamiento del incendio o si se conoce como es el comportamiento del incendio, se podría, con un cierto grado de acierto, determinar cuales son las condiciones ambientales que reproducen este comportamiento. Pero en la práctica, esto no siempre ocurre así, una observación muy importante en este punto es que la segunda opción presenta una notable ventaja cuando se utilizan simuladores de fuego poco precisos, esto se debe a que los métodos determinan como condiciones los parámetros en el contexto del simulador, pero no necesariamente éstos tendrán el mismo valor que las condiciones ambientales reales. Por ejemplo, si se miden todas las variables que representan las condiciones ambientales y suponiendo que todas estas mediciones son correctas, al introducir las variables dentro del simulador, puede reproducir un comportamiento diferente al fuego real, las causas pueden ser: problemas con el modelo matemático o la traducción de éstos a un lenguaje informático.

Ambas metodologías están clasificadas dentro de los métodos guiados

por los datos, en donde la etapa de predicción es llevada un paso de tiempo más adelante y se introduce una etapa de ajuste como se ha descrito en los capítulos 2 y 3, lo que posibilita un cierto grado de independencia sobre los parámetros difíciles de obtener, y donde dichos métodos enfocan sus esfuerzos. Una de las mayores ventajas prácticas de la combinación, se debe a que los dos métodos fueron implementados con el mismo simulador como semilla, el fireLib, descrito en el capítulo 2. Y, como consecuencia se tiene una variedad de factores positivos, tales como:

- Al utilizar el mismo simulador se trabaja con la misma cantidad y tipo de parámetros de entrada, lo que en consecuencia produce que la composición de un individuo en el método evolutivo sea igual a la composición de un escenario en el método estadístico.
- La implementación de los métodos utilizan el mismo paradigma de programación paralela el Master/Worker, con librerías de comunicación MPI, lo que hace que la encapsulación de un método dentro del otro, en términos de codificación algorítmica, represente una tarea menos compleja de implementar.
- Las librerías de datos utilizadas son las mismas para ambos métodos, también la función de calidad descrita en el capítulo 3, por la ecuación 3.2 permite unificar la función objetivo.
- El entorno Hardware/Software de la implementación experimental de los métodos es el mismo, esto permite elaborar medidas comparativas entre los diversos métodos.

4.2. Propuesta - Integación del Método Estadístico con el Evolutivo

Con esta nueva metodología se efectúa un cruce de métodos que busca resaltar las características fuertes de ambos y reducir el impacto de aquellas que debilitan a cada método. Por lo tanto, se obtiene una nueva clase de método en donde se pretende alcanzar las siguientes mejoras:

- Mejoras en el Método Evolutivo:
 - Los algoritmos evolutivos basan su propuesta en la teoría de Darwin, donde la ley fundamental es la sobrevivencia del más apto, pero también es cierto que en la naturaleza existe cooperación entre los individuos y con esta propuesta se quiere aplicar en cierta forma un marco de asociación entre los individuos para conseguir un objetivo, partiendo de la premisa que una población es más fuerte que un individuo.

- Al convertir el método de solución única a solución múltiple, se busca maximizar la capacidad de adaptabilidad, entonces, prevenir modificaciones en las condiciones del incendio en el siguiente instante de tiempo, reduciendo el impacto negativo de cambios bruscos en el clima o la composición de la vegetación.
 - En los casos en donde el algoritmo evolutivo no converge en una solución óptima, con la integración de los individuos se podría reproducir un mapa simulado más semejante al fuego real y conseguir de esta forma un mejor ajuste.
- Mejoras en el Método Estadístico
- Reducir el tiempo de simulación, optimizando los parámetros a un número n de posibilidades, con ello se reduce el espacio de búsqueda, y a su vez, el número de simulaciones necesarias para elaborar la predicción.
 - Adaptar el método con miras a construir una aplicación de tiempo real y elaborar predicciones con un margen de tiempo aceptable.
 - Conseguir una independencia hasta cierto grado de la capacidad del entorno, donde se ejecuta la simulación, eliminando con la optimización la carga innecesaria de cómputo para conjuntos de parámetros inservibles.

Hasta este momento, se ha mencionado cómo mejorar ambos métodos, pero no se ha expuesto el aspecto fundamental de la integración. Es decir, que el problema principal en la predicción de incendios forestales son las diferencias entre el mapa simulado y el fuego real. Ambas metodologías elaboran predicciones con márgenes de error aceptables, pero existen situaciones en donde las simulaciones son de muy baja calidad. Y es aquí en donde se ataca el problema principal, para producir simulaciones más realistas. Por ejemplo, tenemos un conjunto de tres simulaciones A, B y C, más un mapa real con el punto inicial del incendio, a partir de aquí se presentan los distintos acercamientos desde la figura 4.1:

- *Considerando sólo el estadístico:* el barrido de búsqueda exhaustiva produce tres simulaciones A, B, y C, en donde se forman conjuntos con distintas posibilidades sobre las zonas que cubren, y al aplicar las operaciones de intersección y unión entre dichos conjuntos se construyen distintas formas de propagación del incendio simulado, tal como muestra la figura 4.1.a. Por lo tanto, al aplicar el cálculo estadístico concluye que,

$$A \cup B \cup C$$

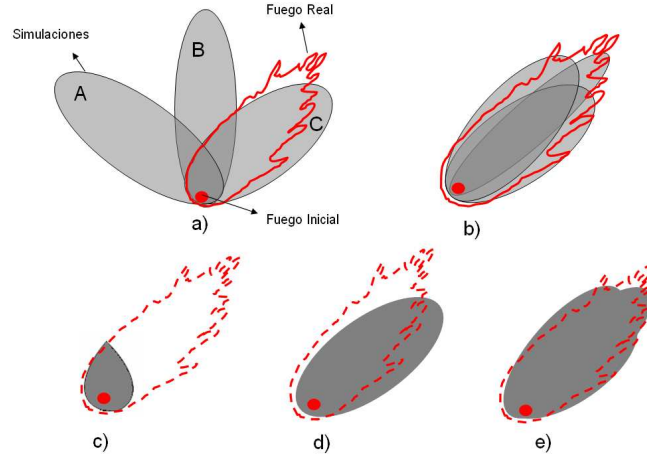


Figura 4.1: Integración de las simulaciones de los distintos métodos

presenta un mapa demasiado extenso en comparación con el fuego real, y,

$$(A \cap B) \cup (B \cap C)$$

presenta un mapa menos extenso que el anterior pero muy diferente al fuego real y por último se observa que,

$$A \cap B \cap C$$

es el que presenta un menor margen de error comparando con los demás conjuntos de posibilidades, pero esto aún es muy diferente al fuego real, tal como muestra el resultado de la figura 4.1.c.

- *Considerando sólo el evolutivo:* ahora en la figura 4.1.a. A, B y C son simulaciones producidas por 3 individuos de una población inicial, los que son evolucionados hasta que dichos individuos produzcan tres simulaciones muy cercanas al mapa real, tal como muestra la figura 4.1.b. Entonces, el algoritmo genético encuentra un individuo que produce un mapa que se acerca más a la forma del incendio y este individuo es el seleccionado para producir el mapa que se representa en la figura 4.1.d. Entonces, puede considerarse que el mapa simulado es mejor que el mapa 4.1.c. producido por el estadístico pero aún presenta diferencias con el mapa real.
- *Considerando la integración:* aplicando primero el método evolutivo tenemos que la figura 4.1.a. son los mapas simulados por individuos

de una población inicial y son evolucionados hasta producir los mapas según la figura 4.1.b. y en este punto se empieza a utilizar el método estadístico, que aplica las operaciones sobre los conjuntos de probabilidades, y en consecuencia, surge como resultado la integración de todos los mapas simulados, ver figura 4.1.e. Se observa que el método estadístico mejora notablemente gracias a que no existen mapas que perjudiquen excesivamente la simulación como en el caso mostrado en la figura 4.1.a.

4.3. Descripción del Método

La idea básica de esta propuesta es que el método evolutivo permita optimizar el conjunto de posibilidades estudiadas por el método estadístico, dirigiendo su espacio de búsqueda a mejores zonas. Es cierto que la fortaleza del método estadístico es su búsqueda exhaustiva y a menor número de posibilidades el método se ve seriamente afectado, pero como se demostró en el capítulo 3.2.4, la búsqueda exhaustiva repercute negativamente en el tiempo de simulación. Entonces, lo que se pretende es reducir la búsqueda exhaustiva de posibilidades y con ayuda del método evolutivo lograr que el efecto de esta estrategia tenga el menor impacto negativo posible. A esta estrategia se le denomina *reducción inteligente del espacio de búsqueda de probabilidades*.

El nuevo método es inicializado igual que el método Evolutivo e introduce dos tipos de algoritmos en la estructura, que se presenta en la figura 4.2.

El algoritmo genético (AG) que incluye toda la funcionalidad presen-

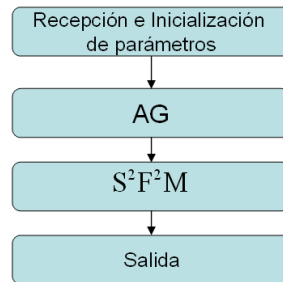


Figura 4.2: Estructura de *SAPIFE*³

tada en el capítulo 3.1, con todas las operaciones (Selección, Cruzamiento y Mutación), pero con una importante modificación, el resultado final ya no será un solo individuo sino toda una población. Consecuentemente, el algoritmo no se detiene cuando encuentra un individuo que produce una simulación con un *error* = 0, y esta población puede tener un número *n* según la configuración inicial. El otro algoritmo introducido es *S²F²M* con

un cambio también muy significativo, la creación de escenarios establecida por la ecuación 3.4 es desechada y suplantada por el resultado del AG. Por lo tanto, el número de escenarios será igual al número de individuos configurado en el algoritmo evolutivo. Desde este punto el método funciona exactamente igual que el estadístico, explicado en el capítulo 3.2.1.

A esta integración de métodos se denomina Sistema Adaptativo para la Predicción de Incendios Forestales basados en Estrategias Estadístico-Evolutivas (*SAPIFE*³).

Por último, se obtiene una salida gráfica que se construye a partir de una matriz de probabilidades, utilizando librerías gráficas de gnuplot [28], de esta forma los ejes X e Y representan la división del terreno en celdas y el eje Z es la probabilidad de ignición de cada celda. Ver figura 4.3.

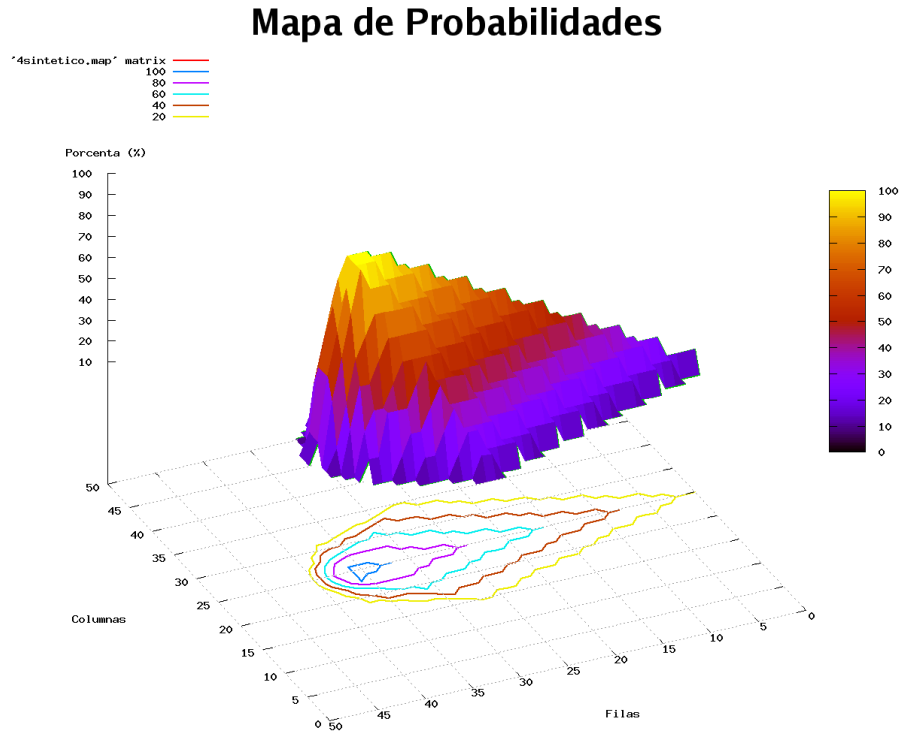


Figura 4.3: Vista en tres dimensiones de la salida de *SAPIFE*³

4.4. Implementación del Método

La idea básica de observar, luego predecir, es la base de ambos métodos utilizados para la integración. Por lo tanto, también es la base de esta metodología. Ahora, considerando la figura 4.4 se define a t_0 como el inicio de todo el método y FR_{t_0} como el frente de fuego en ese instante. Durante el desarrollo del incendio se construye la población inicial con la fórmula propuesta en el capítulo 3, ecuación 3.1.

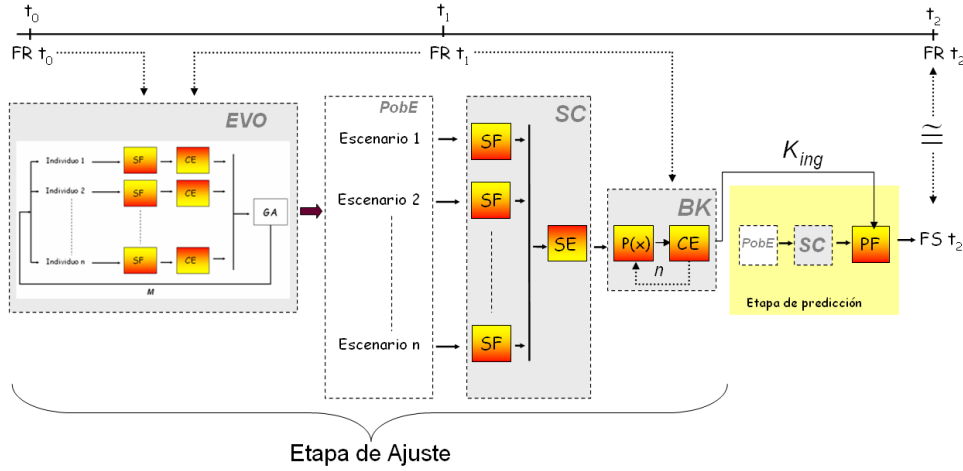


Figura 4.4: Esquema de Implementación de *SAPIFE*³

Una vez que el incendio alcance el instante t_1 se introduce este nuevo frente de fuego al que se lo denomina FR_{t_1} y con él se calcula la función objetivo en el módulo **CE** (Cálculo de Error) y se empieza a construir la población de escenarios **PobE**. Una vez concluidas todas las iteraciones se llega a una población de escenarios evolucionada M veces, hasta este momento el método ha trabajado en el dominio del método evolutivo **EVO** y a partir de aquí, el método empieza a trabajar en el dominio del método estadístico, en donde cada individuo de la población de escenarios es sometida a la función integración dentro del módulo **SC** (Simulación Combinatoria) y donde se construye la matriz de probabilidades de t_0 a t_1 .

El siguiente paso es enviar esta matriz al módulo de búsqueda del factor K_{ign} que reproduzca el valor de probabilidad que mejor se adapte al FR_{t_1} , denominado **BK**. Una vez efectuado todo este proceso se tiene como resultado el K_{ign} que es enviado a la etapa de predicción, en donde primero se construye el mapa de probabilidades de t_1 a t_2 en el módulo **SC** a partir del **PobE** de la etapa anterior. Entonces, el mecanismo de predicción de fuego (**PF**) realiza el corte en la altura del cono, explicado en el capítulo 3.2, figura

3.14, de acuerdo con el K_{ign} recibido de la etapa anterior. Con esto se construye el mapa simulado para la predicción del periodo de tiempo de t_1 a t_2 . Un cálculo adicional es insertado para saber que tan buena fue la predicción, entonces, una vez que se llega al tiempo t_2 se puede comparar con el FR_{t_2} , y así obtener una medida de calidad.

4.5. Implementación en un Esquema Paralelo Master/Worker

El tiempo en esta nueva metodología es una cuestión de vital importancia, debemos utilizar estrategias que disminuyan el tiempo de simulación. Por lo tanto, se debe apelar a la programación paralela. Generalmente los programas se dividen en secuenciales o paralelos, un programa secuencial se caracteriza por no depender de la velocidad de ejecución y de producir el mismo resultado para un mismo conjunto de datos de entrada. En un programa paralelo las actividades que lo constituyen están relativamente superpuestas en el tiempo. Esto significa que una operación puede ser iniciada en función de la ocurrencia de algún evento, antes del término de la operación que estaba ejecutándose anteriormente.

En el capítulo 1 se plantearon algunos paradigmas de programación paralela, de estas alternativas la decisión fue implementar la aplicación utilizando el modelo Master/Worker, con el fin de dar un primer acercamiento de la potencialidad de la metodología expuesta en este trabajo. Básicamente, este modelo consiste en dos entidades: el master y múltiples workers, el master es el encargado de dividir el problema en tareas y distribuir estas tareas a un grupo de workers, una vez que éstos terminan, el master debe recoger los resultados parciales para luego producir el resultado final. La comunicación entre las entidades es por medio de pasos de mensajes utilizando librerías de comunicación MPI. La implementación de *SAPIFE*³ con este modelo puede ser representado lógicamente como una estructura *parbegin-parend* de dos etapas secuenciales (ver figura 4.5). Cada tarea en este esquema es expuesto como S que dentro de la metodología constituye una simulación de un escenario (en este punto cabe observar que cuando se refiere a escenario es también aplicable al individuo), y los nodos del grafo *parbegin-parend* corresponden al master. En el diagrama se puede constatar las características principales de este modelo, donde en el punto *parend* es necesario una fuerte sincronización y todas las simulaciones son integradas, obsérvese que no existe comunicación entre las tareas S sólo desde el master al worker y viceversa. Una vez terminado el proceso en el dominio del evolutivo, comienza la segunda fase, donde el master envía los escenarios que son los resultados de la evolución a los workers, que devuelven una matriz booleana y una vez que todos los escenarios fueron procesados el master construye el

mapa de probabilidades.

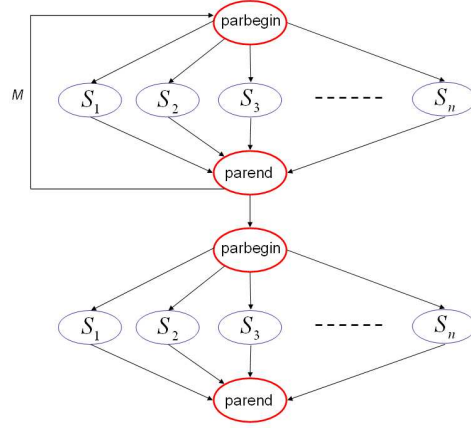


Figura 4.5: Estructura parbegin-parend de *SAPIFE*³

En otras palabras, el master envía un individuo o escenario al worker y éste devuelve el resultado de la simulación, entonces el master vuelve a enviar otra simulación al worker y así sucesivamente, hasta que todas las simulaciones hayan acabado. En este punto se puede constatar que el número de tareas ejecutadas en paralelo depende exclusivamente de la cantidad de nodos en donde se ejecute la aplicación.

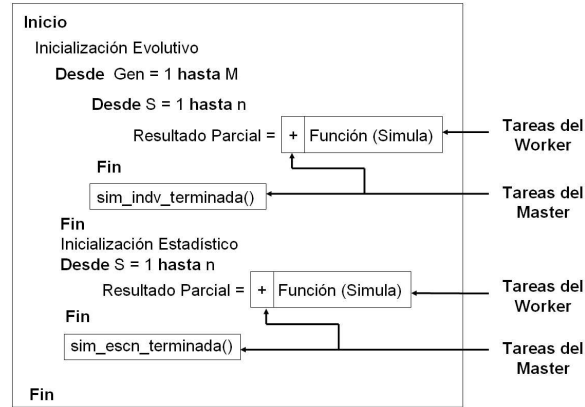


Figura 4.6: Pseudocódigo Simple de la Implementación de *SAPIFE*³ en el modelo Master/Worker

En la figura 4.6 se describe el pseudocódigo de esta implementación, en donde primero es ejecutado el bucle correspondiente al método evolutivo y

donde todo el proceso de asignación, coordinación y compilación es responsabilidad del master, que envía el conjunto de parámetros al worker para que ejecute la función (simula) que devuelve el individuo con su valor de error, toda esta operación es ejecutada M veces que corresponde al número de generaciones. Una vez terminada todas las generaciones se pasa al siguiente bucle, donde al igual que el primero, toda la responsabilidad de asignación, coordinación y compilación de los resultados es tarea del master y donde cada worker recibe los escenarios y devuelve una matriz booleana que representa la celdas alcanzadas por el fuego. El master es el encargado de compilar esta información para generar el mapa de probabilidades y realizar el ajuste y la predicción. [29]

Capítulo 5

Resultados Experimentales

Los incendios dependen de factores naturales o físicos y de factores humanos o sociales. Entre esos factores están la meteorología, la vegetación, el tiempo y la topografía. Y son estas condiciones las que pueden ser representadas dentro de un laboratorio, en el caso particular de este trabajo el laboratorio es la computadora. Esto permite generar incendios ficticios en un entorno virtual y manipular las condiciones a medida de los requerimientos de la propia experimentación, utilizando como herramientas los simuladores de fuego. Pero la ciencia computacional, detallado en el capítulo 1, permite interactuar con distintas disciplinas, por lo que el marco de acción de estos experimentos se extiende un poco más allá para contrastar las hipótesis y teorías planteadas con incendios de quemas reales realizadas en campos de experimentación preparados para el efecto.

Antes de comenzar el detalle propio de la configuración de los experimentos se definen los objetivos y el alcance de las pruebas realizadas. El objetivo es comprobar si la propuesta de integración de los métodos consigue mejorar los resultados obtenidos por las otras dos metodologías. Para ello, el foco de observación es el comportamiento del fuego, no se tiene en cuenta la causa o los factores humanos o sociales. Se considera al fuego como una entidad física, con dimensiones en un espacio bidimensional lo que facilita la comparación de formas entre los resultados de la simulación y el fuego real.

Para evaluar los resultados de *SAPIFE*³, fue necesario compararlo con los demás métodos, Evolutivo [25] y Estadístico S^2F^2M [27], que fueron evaluados independientemente en otros trabajos. Consecuentemente, esto representa una contribución notable a la hora de validar los resultados. Pero una observación más que importante se refiere a que los experimentos realizados en S^2F^2M pueden presentar algunas diferencias, debido a que la unidad de medida o función objetivo consistía en un fitness (ver ecuación 5.1) con un rango 0 a 1, siendo el *fitness* = 0 el resultado más pobre y *fitness* = 1 el mejor resultado posible. Esta medida no refleja completamente la calidad de la simulación, tal como muestra el ejemplo de la figura 5.1, en donde ca-

da mapa del ejemplo contiene un total de 9 celdas y se divide en regiones en donde la línea naranja encierra las celdas quemadas por el fuego real produciendo de esta manera las siguientes divisiones: la zona cuadrículada corresponde a las celdas de inicio del fuego (*CeldasInicio*), las celdas de intersección (*Celdas*∩) están compuestas por las celdas iniciales más las celdas verdes que representan las celdas quemadas en la simulación y el fuego real, las celdas de la unión (*Celdas*∪) para el primer mapa es la suma de las celdas iniciales más las celdas de la intersección y las celdas rojas, que son las quemadas por el fuego real y no por la simulación, para el segundo caso, las celdas de la unión están compuestas por las celdas iniciales, las celdas de la intersección y las celdas amarillas que son las quemadas por el simulador y no por el fuego real. Vemos que el primer mapa, tiene menor diferencia que el segundo, pero aplicando la fórmula de fitness (ver ecuación 5.1) el resultado final es el mismo, y aplicando la fórmula de error (ver capítulo 3, ecuación 3.2) sí se ven reflejadas numéricamente la diferencias en la calidad de la simulación como se demuestra gráficamente.

$$Fitness = \frac{Celdas \cap - CeldasInicio}{Celdas \cup - CeldasInicio} \quad (5.1)$$

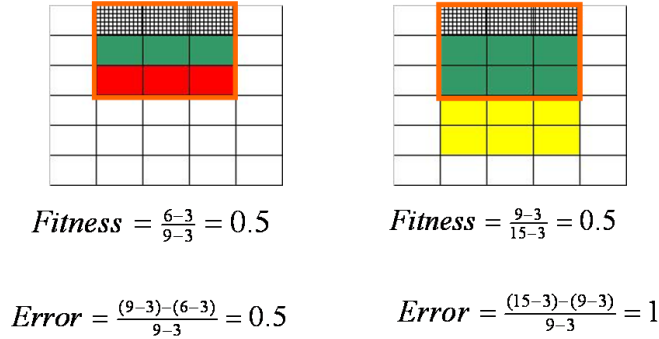


Figura 5.1: Función Fitness

Esto puede afectar de manera significativa el comportamiento y el resultado de S^2F^2M , y para evaluar correctamente los resultados de este método fue necesario incorporar la función de error y a partir de allí, comparar con $SAPIFE^3$ y el Evolutivo.

La configuración de $SAPIFE^3$, contiene en gran medida las variables que corresponden al método evolutivo, para el que se realizaron varias pruebas de tal forma a configurar y obtener los mejores resultados. Para dicha configuración, se tomaron como referencia los experimentos realizados por Mónica

Denham [25], con algunas variantes teniendo en cuenta que el potencial del método estadístico es el barrido exhaustivo, se buscó representar un número de individuos que abarque una muestra considerable de las posibilidades sin perjudicar el tiempo de simulación, además, a pesar de contar con la posibilidad de guiar la búsqueda, esta opción ya no fue incluida dentro de estos experimentos debido a que los resultados preliminares demostraron que guiar la búsqueda produce individuos muy semejantes, lo que afecta negativamente al método integrado. Entonces, se optó por las siguientes configuraciones:

- Población: se utiliza una población inicial de 500 individuos, donde los valores para los parámetros son aleatorios con distribución uniforme en cada rango válido de variación. Esto intenta garantizar variabilidad en la población.
- Cantidad de generaciones: se pudo observar que en las primeras generaciones el algoritmo genético encuentra buenos individuos y que el error encontrado se suele estabilizar antes de las primeras 5 iteraciones. Por esto, se utilizan 5 iteraciones del algoritmo genético como se mostró en el capítulo 3, figura 3.7.
- Probabilidad de mutación: se utiliza una probabilidad de 1 % para realizar la operación de mutación. Como en la naturaleza, la mutación ocurre con muy bajas probabilidades. Esto es para evitar que el algoritmo genético se convierta en una búsqueda aleatoria.

Cada experimento muestra dos tipos de resultados que corresponden a las etapas de ajuste y predicción. Y debido a que cada simulación tiene un factor aleatorio que corresponde a la generación de la población inicial, todos los resultados son promedios de 5 poblaciones distintas. Cada experimento es dividido en periodos cortos de tiempo y dependiendo del tiempo inicial y final se pueden obtener distintos números de pasos para cada experimento. Esto permite ver como los métodos van adaptando el comportamiento de acuerdo al paso de ajuste, que es el periodo en donde se compara el comportamiento del incendio. En el entorno del experimento, es considerado fuego real el mapa de fuego utilizado como base sin importar su origen genérico, la computadora o una quema real.

En el caso del fuego real originado por computadora se lo denomina “Mapa Sintético”, que se genera introduciendo valores al simulador fireLib que es encargado de elaborar el mapa de propagación, luego estos parámetros son desechados y sólo el mapa es utilizado como incendio real.

5.1. Mapas Sintéticos

El diseño de los experimentos para mapas sintéticos se dividió en dos tipos de condiciones del entorno medio-ambiental, el primero consiste en un ma-

pa de comportamiento estable de las condiciones durante la propagación del fuego, en el segundo se han insertado variaciones sobre las condiciones medios-ambientales, actuando sobre aquellas variables dinámicas como viento y humedad, de manera a que la variación de estos parámetros provoquen un comportamiento irregular de la propagación del incendio, con el fin de medir cómo los métodos reaccionan a esta situación muy frecuente en el entorno de un incendio real.

5.1.1. Sintético I

Este mapa tiene una dimensión de 70 metros de alto y 90 metros de ancho, cada porción del terreno fue dividida en celdas de 1 metro de alto y 1 metro de largo. La pendiente tiene una inclinación de 18 grados y el modelo de vegetación corresponde a arbustos pequeños con ramas, definido como modelo 7 dentro de las librerías de datos del fireLib, descrito en el capítulo 2. Un primer paso a considerar es la conversión de la variable continua de tiempo al tipo discreta, dividiendo los pasos en periodos o rangos de tiempo de dos minutos, de esta forma los métodos, efectúan un punto de control cada 2 minutos observando la forma que tiene el incendio en ese instante y predicen el comportamiento para el siguiente periodo.

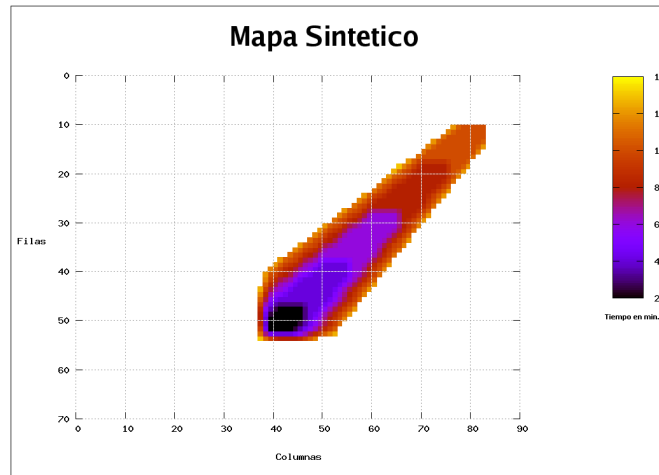


Figura 5.2: Sintético I

En la figura 5.2, se observa la forma que va adquiriendo el fuego, y puede notarse la forma elíptica como consecuencia del modelo físico implementado en el simulador. El periodo de observación del incendio, comprende de 2 a 14 minutos, lo que permite dividir el incendio en 5 pasos de ajuste y 5

pasos de predicción. También se puede notar el comportamiento regular del incendio, debido a que las condiciones y el entorno del incendio se consideran uniformes durante toda la propagación.

El siguiente paso, consiste en analizar los resultados conseguidos, que se muestra en la figura 5.3, en donde el eje y representa el error obtenido de un promedio de 5 casos, para cada uno de ellos se utilizaron diferentes poblaciones aleatorias, y el eje x presenta los intervalos de tiempo de cada paso. Consecuentemente, tenemos que las dos primeras barras se encuentran sobre el periodo 2-4-6 que expresa la unión del tiempo de ajuste y predicción, por ejemplo, la etapa de ajuste va de 2 a 4 y la etapa de predicción va de 4 a 6 minutos. Los resultados obtenidos nos muestra una regularidad en todo el proceso, el error gira en torno al 20 %. Los resultados muestran una correspondencia entre el error en la etapa de ajuste y la predicción. Entonces, esto nos permite concluir que en un entorno uniforme con condiciones estables, los resultados de la etapa de ajuste condicionan la calidad de la predicción. En otras palabras, a un mejor ajuste una mejor predicción.

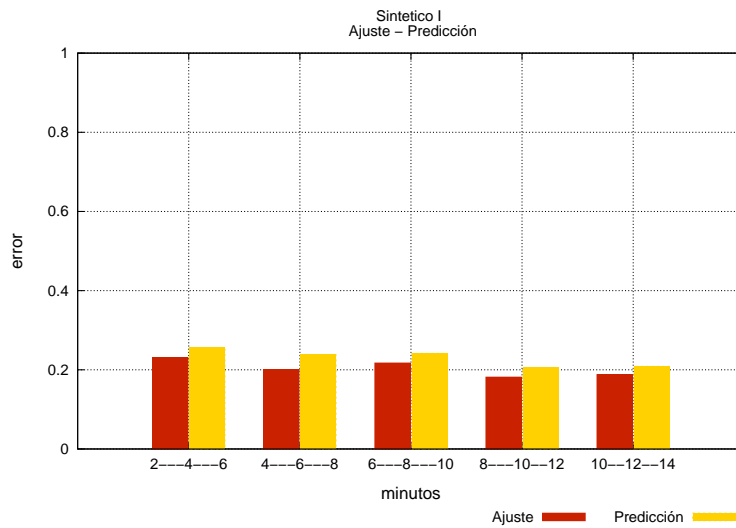


Figura 5.3: Mapa sintético I, errores medidos en la etapa de ajuste y predicción

Una observación importante se refiere al Algoritmo Genético utilizado por el método, en donde algunos individuos evaluados, producto de las operaciones evolutivas, en algunos casos presentaban un valor de error más bajo que el resultado final de todo el método, la diferencia en promedio representaba un 3 % más de error que los mejores individuos de cada paso de tiempo. Esto se debe a que la integración de algunos individuos, no tan buenos, disminuyen la calidad de la simulación, como se explicó en el capítulo 3.2.4, pero el

efecto negativo de estos escenarios considerando al método estadístico, se ha reducido notablemente gracias a la optimización.

El siguiente paso, fue elaborar la experimentación de tal forma a comparar los tres métodos y corroborar los supuestos en el experimento anterior. Consecuentemente, las pruebas incluyen la configuración y ejecución de S^2F^2M (ver cuadro 5.1), en un cluster de 32 nodos. Además, el Algoritmo Evolutivo fue configurado según muestra el cuadro 5.2 y fue ejecutado en un cluster de 8 nodos.

Parámetro	Límite Inferior	Límite Superior	Paso
Model	0	13	1
WindSpeed	0,67	12,63	0,67
WindDir	0	360	5
Slope	18	18	1
Aspect	180	180	1
M1	0,1	0,3	1,99
M10	0,1	0,3	1,99
M100	0,1	0,5	1,99
Mherb	0,1	0,1	1

Cuadro 5.1: Configuración de Parámetros de S^2F^2M

Parámetro	Valor
Model	7
sizePopulation	500
numGenerations	5
elitism	20
crossoverProbability	0.2
mutationProbability	0.01
guidedMutation	no
guidedEllitism	no
slope	18
aspect	180

Cuadro 5.2: Configuración de Parámetros del Algoritmo Genético

Los resultados de las comparativas están basados en la etapa de ajuste, tal como muestra la figura 5.4, debido a que el resultado de la predicción es un contexto uniforme, con comportamiento estable de las condiciones medio-ambientales, la predicción no debería presentar valores de errores

muy diferentes a los conseguidos en la etapa anterior. Se puede observar que *SAPIFE*³ llega a producir errores muy similares al evolutivo y reafirma la suposición planteada con el experimento anterior, en donde algunos individuos del propio método superan al resultado final. Esto es fácil de comprender, porque el método elabora una integración de varios individuos entre los que puede existir algunos malos que perjudiquen la forma final de la propagación. Sin embargo, el evolutivo escoge siempre al mejor y considerando el entorno uniforme es razonable que tenga un mejor resultado.

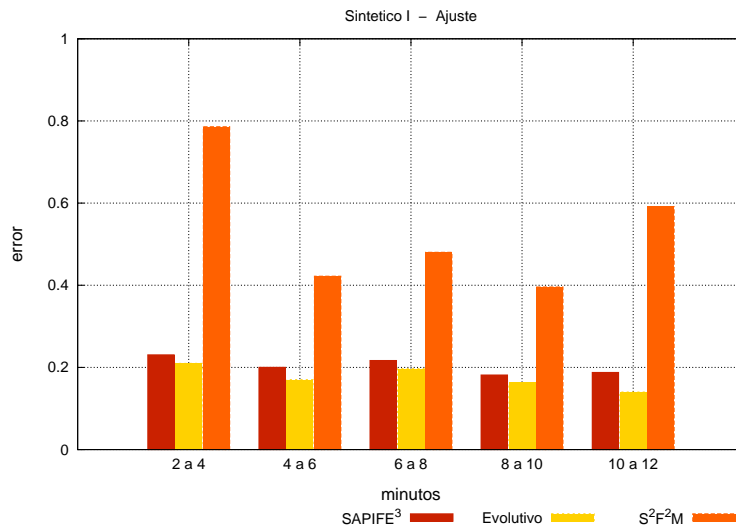


Figura 5.4: Mapa sintético I, comparativa con los tres métodos

El Evolutivo obtiene los mejores resultados con respecto a los otros dos métodos, llegando a tener errores por debajo del 20 % durante los pasos de simulación, claramente se puede observar que el evolutivo ha encontrado valores de parámetros buenos, pero no los óptimos, aunque se efectúe varias evoluciones no se consigue mejorar los resultados de una manera considerable, esto se puede constatar en el último paso de simulación, considerando que al llegar a este paso de simulación el número de evoluciones alcanzada es de 25. Por último, se puede observar que el peor rendimiento lo presenta *S²F²M*, esto se debe a que el incendio es de una característica similar al caso presentado entre las desventajas del método estadístico en el capítulo 3.2.4, en donde existen simulaciones que perjudican excesivamente la predicción. Pero también, cabe destacar que el comportamiento del estadístico una vez que se ajusta, aunque el valor de error es considerablemente alto, se mantiene estable durante el resto de los pasos de simulación.

5.1.2. Sintético II

El diseño de los experimentos para este mapa, tiene en consideración un terreno más largo que el anterior con medidas de 120 metros de alto y 80 metros de ancho, dividiendo las celdas con una dimensión de 1 metro de largo y alto. La pendiente tiene una inclinación de 18 grados y el modelo de vegetación es el número 7, similar al utilizado para el mapa anterior. El tiempo esta dividido en periodos de 2 minutos.

La principal diferencia con el experimento anterior, es la variación de las condiciones ambientales en un punto específico en el tiempo. La finalidad es recrear un suceso muy frecuente en un entorno real, donde las variables dinámicas se comportan de una manera totalmente arbitraria. Para el efecto, los parámetros elegidos para la alteración son la humedad de la vegetación y el viento. La razón de modificar estos valores de parámetros se debe a que en la realidad el primero es muy difícil de medir en tiempo real, además el incendio modifica esta variable de acuerdo a su intensidad, y el segundo es uno de los factores más dinámicos e influyentes en el comportamiento del incendio. Por lo tanto, esto produce la división del mapa en dos periodos largos. Considerando al primer periodo de tiempo de 2 a 8 minutos, de clima y humedad estable, así como muestra la figura 5.5.

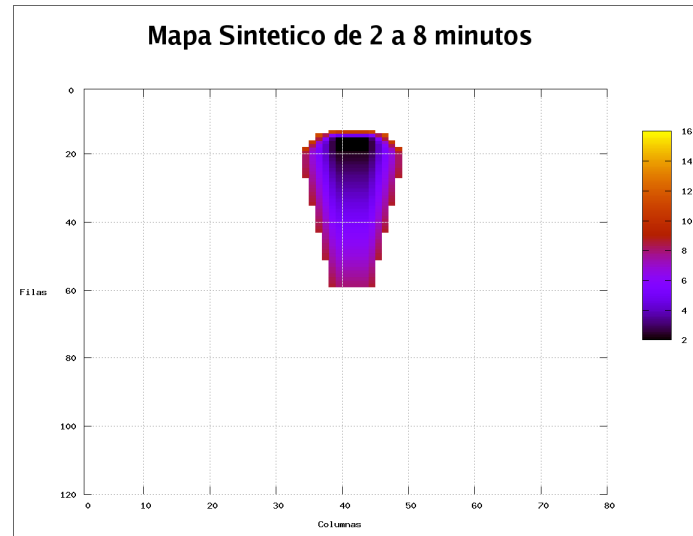


Figura 5.5: Primer Periodo del Sintético II

En el minuto 8 de la propagación se altera el valor de la humedad, reduciéndola de un 35 % a un 5 %, con respecto al viento, se aumenta la velocidad de 5 a 10 millas por hora y se rota su dirección de 270 a 180 grados, los demás parámetros que corresponden a la topografía y la composición del terreno se mantienen uniformes. Todo esto produce el mapa de la figura 5.6.

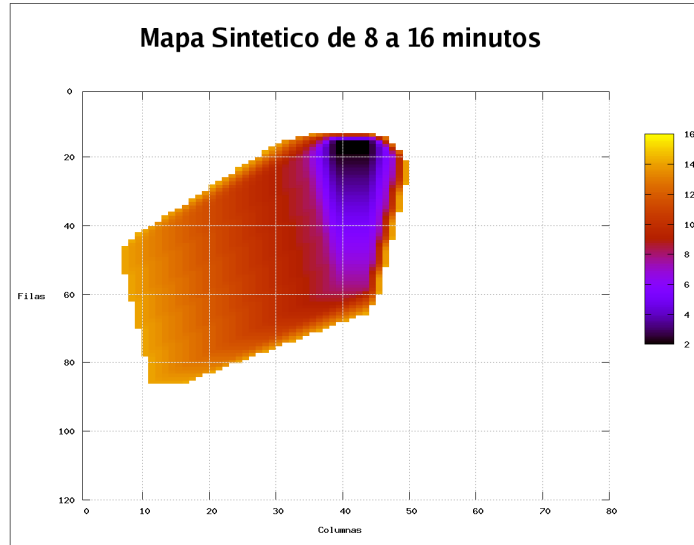


Figura 5.6: Segundo Periodo del Sintético II

Este segundo mapa es el que corresponde al comportamiento final del incendio desde el minuto 2 al 16, observamos como la forma del incendio se vé sensiblemente alterada, este tipo de situaciones son unas de las causas de malos resultados en los métodos clásicos de predicción tratado en el capítulo 2.1. Y es este entorno el que necesita de la adaptación dinámica.

Esta claro, que la experimentación con mapas sintéticos nos permite controlar las variables y sus efectos, pero una observación más que importante es que el sistema nunca tiene conocimiento de cuáles fueron los factores alterados, lo único relevante para el sistema en este caso es la forma geométrica que el fuego va adquiriendo.

La configuración de *SAPIFE*³, es similar al caso anterior del mapa sintético I, con 500 individuos, y esta configuración no es modificada durante todo el proceso de simulación, de esta forma el objetivo es observar como responde el sistema a estos cambios en el comportamiento del incendio. Para medir los resultados de esta experimentación se presenta la figura 5.7.

Este gráfico presenta en el eje y el error que mide las diferencias entre el mapa simulado y el real, en el eje x los pasos de simulación para cada par de periodos de tiempo, que al igual que el mapa sintético I, esta compuesto de dos periodos de tiempos divididos para el ajuste y la predicción. De esta manera se demuestra que los dos primeros pasos de simulación (2-4-6 y 4-6-8) presentan mucha similitud con las gráficas de la figura 5.3 en donde el método tiene un comportamiento estable, entonces se puede afirmar que la predicción se corresponde con la etapa de ajuste. Y en el tercer paso de sim-

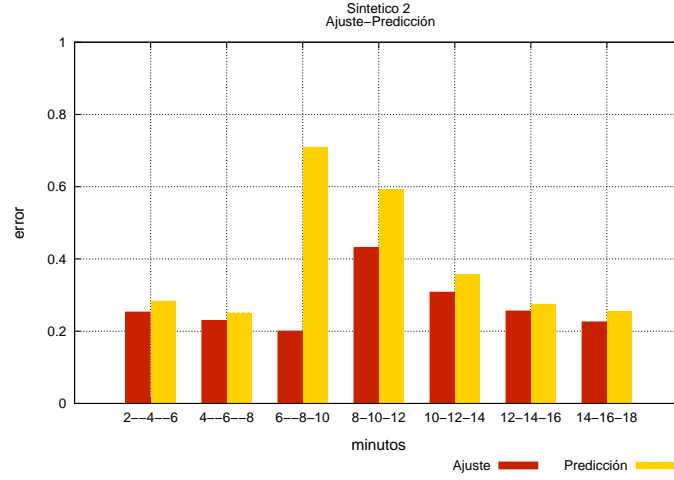


Figura 5.7: Mapa Sintético II, errores medidos en las etapas de ajuste y predicción

ulación (6-8-10), el método aún está estable para los resultados del ajuste, pero es en este punto donde se produce un salto del valor de error como consecuencia del cambio en las condiciones del incendio, que se insertaron artificialmente en este caso. Se puede notar que el ajuste está basado en una situación anterior a la de la predicción, y que las condiciones que regían ese periodo de tiempo, que va de 6 a 8 minutos, se han modificado considerablemente a partir del minuto 8 en adelante. Lo importante de este experimento es que demuestra que el método vuelve a reajustarse dinámicamente en el siguiente paso de tiempo hasta estabilizarse y conseguir de nuevo un nivel aceptable de error.

La adaptación dinámica es una de las grandes ventajas de los métodos guiados por los datos, pero el tiempo es un factor clave, qué tan rápido puede el sistema adaptarse y prevenir estos cambios, y es aquí donde se nota la diferencia entre los métodos para este caso. Entonces, se hace necesario determinar si los resultados son buenos o malos, por lo que se debe establecer una comparativa con los demás métodos, S^2F^2M y el Evolutivo, y analizar de qué manera se absorbe el impacto de un cambio brusco en las condiciones. Para ello, se realizaron experimentos con los tres métodos que incluyen las mismas configuraciones presentadas para el caso anterior, con el mapa sintético I, en donde la configuración de S^2F^2M se observa en el cuadro 5.1 y la configuración del Evolutivo con el cuadro 5.2.

La figura 5.8 muestra los resultados obtenidos por los tres métodos. Se puede observar que los tres primeros pasos de la simulación tienen un comportamiento muy similar al del caso anterior con el mapa sintético I y una vez más se comprueba que, cuando el entorno es uniforme y las condiciones

estables los métodos consiguen un valor de error aceptable para distintas situaciones de incendio. Bajo estas condiciones, el método Evolutivo es relativamente mejor que $SAPIFE^3$ y S^2F^2M , viéndose este último seriamente afectado, debido a que como no es un incendio lineal las simulaciones que producen mapas en distintas direcciones afectan negativamente y elevan el nivel de error (ver capítulo 3.2.4).

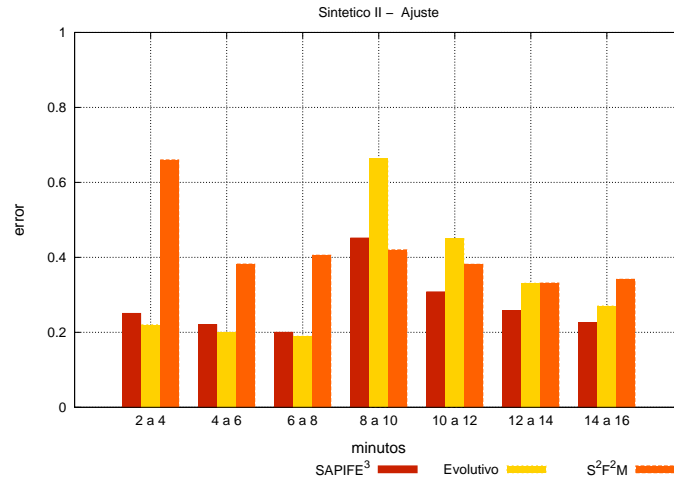


Figura 5.8: Mapa Sintético II, comparativa entre los tres métodos

Entonces, se llega al punto en donde se introdujo las variaciones en las condiciones, el paso de 8 a 10, se observa como los métodos se ven seriamente afectados y el valor de error se dispara considerablemente. Pero aquí, podemos notar dos puntos muy importantes, el primero, se refiere al valor de error de $SAPIFE^3$, sí es cierto que también se dispara, pero es mucho menor error que el producido por el Evolutivo. Y el segundo punto, es lo que se refiere al S^2F^2M que no se ve tan afectado por este cambio brusco de condiciones ambientales, pero como el valor de error era relativamente alto en el ajuste, tampoco se puede decir que este método sea mejor en este caso. $SAPIFE^3$ consigue prácticamente equipararse al método estadístico y rápidamente se recupera en los siguientes paso de simulación. Una observación interesante, es que el incendio empieza a tener una forma de avance más lineal, es por ello, que los ajustes de S^2F^2M empiezan a mejorar considerablemente. Otro aspecto también relevante es que a partir del minuto 8 el comportamiento de las variables ambientales vuelve a ser estable, entonces los métodos se estabilizan y consiguen valores de error aceptables, una vez que consiguen ajustarse al nuevo comportamiento.

En la figura 5.9, se puede apreciar de qué manera afecta el cambio brusco

de las condiciones climáticas en la etapa de predicción, y se observa que el que mejor valor de ajuste presenta es el Evolutivo, pero en la predicción el error se dispara más que los demás métodos y el que se ve menos afectado es S^2F^2M , en donde su valor de ajuste no es tan diferente al de su predicción. Esto confirma la suposición de que S^2F^2M , esta mejor diseñado para prever cambios, pero su valor de error es muy elevado en comparación de los otros dos métodos en la etapa de ajuste.

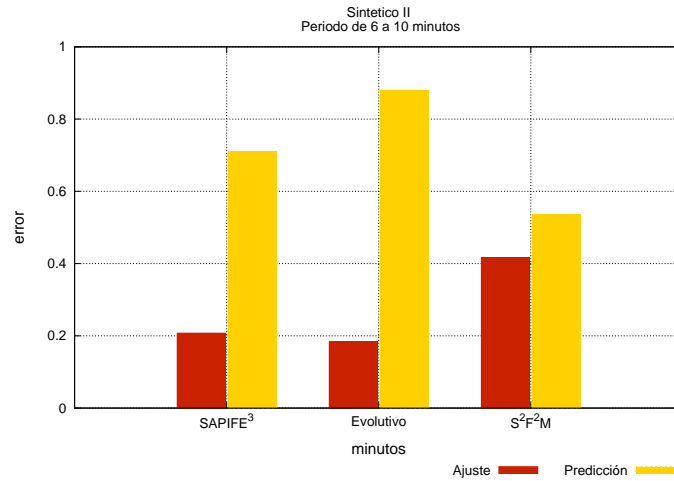


Figura 5.9: Impacto de los cambios bruscos en los tres métodos

Con este experimento, se demostró que $SAPIFE^3$, producto de la integración, ha conseguido rescatar las fortalezas de ambos, este contexto se refiere a que en la etapa de ajuste el método es considerablemente mejor que S^2F^2M y no tan diferente al Evolutivo que obtiene el mejor resultado. Al mismo tiempo, cuando se presenta un cambio brusco, el método ha conseguido un mejor resultado que el Evolutivo, a pesar de tener una diferencia con el estadístico, que consigue el valor de error más bajo. Esto es lo que da pie a la integración, un buen ajuste y menos susceptible a cambios bruscos.

5.2. Mapas Reales

Los mapas de fuego reales, tratados en estos experimentos corresponden a quemas controladas, también conocidas como fuegos prescritos, con la finalidad de elaborar estudios sobre el comportamiento del fuego. El área de estudio está ubicada en el centro de Portugal (40°15'N, 8°10'O) en una ladera de Serra da Lousã, a una altitud entre 800 y 950 m sobre el nivel del mar. Las quemas fueron realizadas en el marco del proyecto SPREAD [30] (los experimentos de campo Gestosa 2002 se iniciaron en el año 1998 y han con-

tinuado hasta la actualidad, involucrando diversos grupos de investigación en el ámbito de distintos proyectos de la Unión Europea). Para mantener la seguridad de las quemas y proveer una buena organización, el área fue dividida en plots con dimensiones regulares, separados mediante cortafuegos para evitar la propagación de éste fuera de los límites deseados. Estas quemas fueron realizadas en tierras del Servicio de Forestación, en la zona de Gestosa [27].

La principal diferencia con los mapas sintéticos es que en este tipo de experimentos ya no se tiene control sobre las variables medio-ambientales. Además, las condiciones climáticas son muy diferentes a los casos anteriores, donde realmente se presenta una situación mucho más aproximada a los incendios reales, con constantes cambios en la velocidad y dirección del viento. Si bien, el terreno fue seleccionado con una textura homogénea, ciertas partes de la topografía pueden presentar algunas diferencias en la humedad o el tipo de combustible.



Figura 5.10: Vista del área de Gestosa, donde se realizaron las quemas prescriptas

5.2.1. Plot 520

El terreno, en muchos aspectos tiene las mismas características que los casos anteriores, en donde la pendiente es de 18 grados y el modelo de vegetación se considera que es el número 7, la superficie es de 89 metros de ancho por 109 metros de alto, cada celda fue dividida con una extensión de 1 metro de largo y ancho.

La figura 5.11, muestra el progreso del incendio que va de 0 hasta 14 minutos,

esto posibilita dividir el comportamiento del incendio, en 5 pasos de ajuste y 5 de pasos de predicción. La propagación es del tipo lineal, debido a que el frente de fuego presenta una forma de ola, esto en gran medida se debe a que las celdas de inicio del fuego estaban diseminadas como una línea recta. Un aspecto muy importante de este incendio es su comportamiento irregular durante casi toda la propagación, esto supone un gran desafío para los diversos métodos, que deberán ser sometidos a continuos cambios de las variables medio-ambientales, durante todo el incendio.

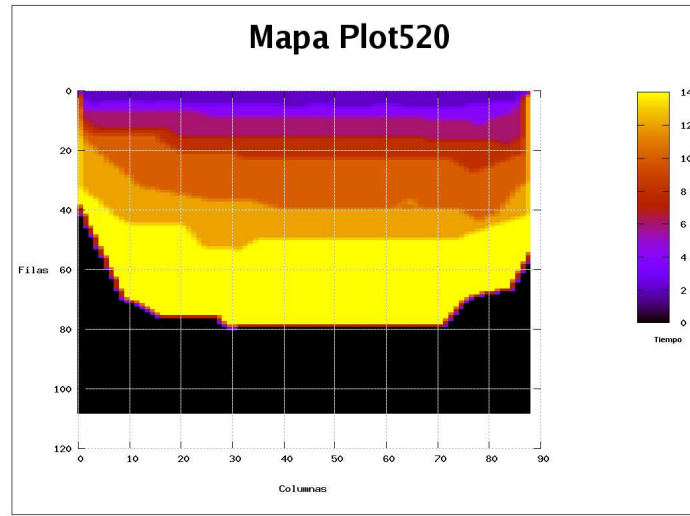


Figura 5.11: Mapa de fuego real, plot 520

Para este caso, al igual que los anteriores, se han elaborado experimentos para comparar *SAPIFE*³ con los demás métodos guiados por los datos, donde las configuraciones del Evolutivo y *S²F²M* son las mismas que las utilizadas para los mapas sintéticos, presentados en los cuadros 5.2 y 5.1. En los casos del Evolutivo y *SAPIFE*³, los resultados son un promedio de 5 poblaciones aleatorias distintas, que significó un experimento para cada una, y en el caso de *S²F²M*, se hicieron 5 repeticiones con la misma configuración, que para cada una el método exploró una cantidad de 28470 escenarios, como resultado de la ecuación 3.4 (ver capítulo 3.2.2, ecuación). En cada paso de tiempo, esto generó una carga de cómputo muy importante, y a pesar de ser ejecutado en un cluster de 32 nodos, el tiempo de simulación era muy superior a los otros dos métodos, mientras los experimentos de *SAPIFE*³ y Evolutivo duraban un promedio de 20 minutos, *S²F²M* invertía en promedio 5 horas.

En la figura 5.12, se muestran los resultados obtenidos de los distintos métodos, en el eje x se presentan los periodos de tiempo del ajuste y en el

eje y los errores obtenidos. Para los dos primeros pasos de tiempo (2 a 4 y 4 a 6), se observa una regularidad entre los tres métodos, donde el evolutivo consigue el mejor ajuste. A partir del tercer periodo (6 a 8), se observa que el evolutivo comienza a obtener valores de error más elevados que el resto, esto tiene su explicación, si se observa el mapa de la propagación figura 5.11, el fuego a partir de ese periodo empieza a comportarse de manera más irregular, esto se supone, es causa de los cambios en las variables medio-ambientales, que a partir de ese momento son más dinámicas variando de un periodo a otro. Esto produce que el fuego en un periodo de tiempo se acelere y en el siguiente se frene.

También, se observa cómo S^2F^2M comienza a obtener el error más bajo, claro está, que las características del fuego lo favorece. Pero aún así, $SAPIFE^3$, se ha ubicado muy cerca de los valores de error obtenidos por este método. En los últimos periodos de tiempo, la dificultad de ajustarse mejor es una tarea más que difícil, debido a que el fuego se comporta de una manera impetuosa, cosa que perjudica considerablemente al Evolutivo, que no consigue un individuo que reproduzca este tipo de propagación, a causa de que su zona de búsqueda, para este periodo esta muy acotada por las constantes evoluciones. Los demás métodos tratan de ajustarse, y en cierta forma, bajo estas condiciones logran un error aceptable, y en caso de que estas condiciones se mantengan constantes durante algún tiempo, probablemente se adaptarán mucho más rápido que el Evolutivo, como se demostró en el caso anterior.

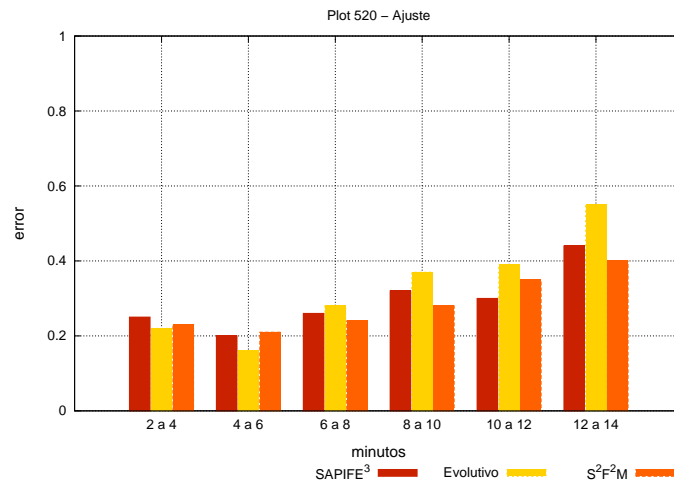


Figura 5.12: Plot 520, errores obtenidos en la etapa de ajuste

La figura 5.13, muestra los resultados de error obtenidos por los distintos

métodos, para la predicción, y vemos una gran irregularidad en los valores, esto se debe, principalmente, a que las condiciones que regían la etapa de ajuste, en la etapa de predicción, se supone eran muy distintas. Esta situación es muy similar a la presentada en el caso del mapa sintético II, donde las condiciones de la etapa de ajuste cambiaron bruscamente en la etapa de predicción, y produjo un salto del nivel de error, pero la principal diferencia, es que en ese caso la variación sólo era en un punto específico del tiempo, y para este caso real, esas variabilidades de las condiciones se producen constantemente, durante toda la propagación del incendio. Pero, a pesar de ello, se observa que *SAPIFE*³ tiene un valor de error más bajo que el Evolutivo, confirmando así los resultados del caso anterior, donde el impacto del cambio de condiciones perjudican menos a *SAPIFE*³ y si bien, *S²F²M* tiene el resultado más bajo de error, hay que considerar que las diferencias son pequeñas y este método ha tardado, en promedio, 15 veces más que *SAPIFE*³.

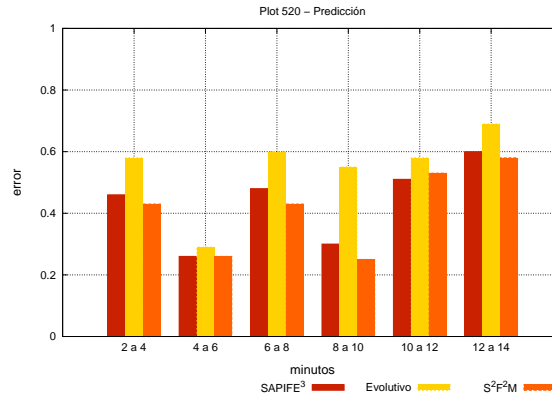


Figura 5.13: Plot 520, errores obtenidos en la etapa de predicción

5.2.2. Plot 751

Considerando, que las variables climatológicas no son de dominio de control de la experimentación, y que se realizaron otras quemas en la misma zona, pero con distintas circunstancias ambientales, con el fin de enriquecer el análisis de rendimiento de los tres métodos, se toma como un caso de estudio más, un segundo mapa real que corresponde al Plot 751.

El terreno tiene menor dimensión que el anterior, 60 metros de ancho y 90 metros de alto, y cada celda mide 1 metro de ancho y 1 metro de largo. La pendiente del terreno es de 6 grados, muy inferior al caso anterior. Una acotación importante es que la pendiente, para este caso no representa una gran influencia sobre el comportamiento del incendio. El modelo de

vegetación supuesto, es también el número 7, al igual que el resto de los casos.

La figura 5.14, se observa que la propagación del incendio, tuvo una duración aproximada de 0 a 12 minutos, lo que permite dividir el tiempo en periodos de 2 minutos, con 4 pasos de ajuste y 4 pasos de predicción. A simple vista, se puede notar que el comportamiento del incendio es mucho más regular que el anterior. Esto puede ser causa de un clima muy estable y un terreno más homogéneo que el plot 520.

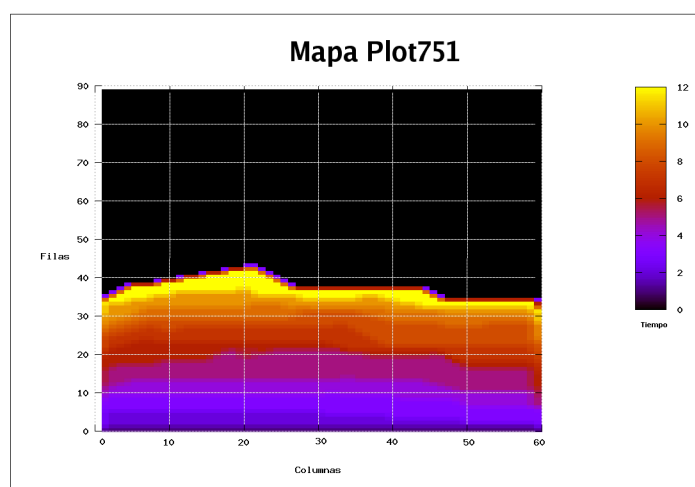


Figura 5.14: Mapa de fuego real, plot 751

Para realizar la comparativa entre los tres métodos, se sigue la misma configuración de los experimentos para los casos anteriores, con una variación en los cuadros 5.2 y 5.1, donde el valor de pendiente (SLOPE), fue ajustado para este caso a 6 grados. Los resultados en el paso de ajuste se puede examinar en la figura 5.15, donde el eje x presenta los periodos de tiempos y el eje y el promedio de error. Los tres métodos consiguen muy buenos resultados, y en algunos casos están por debajo del 20 % de error. Esto en gran medida se debe a lo homogéneo del terreno y las condiciones medio-ambientales, tal como se demostró en el primer caso con el mapa sintético I. Además, el incendio presenta un comportamiento lineal, uno de los tipos de incendios donde mejor se desempeñan estos métodos.

El Evolutivo consigue acercarse a una muy buena zona de búsqueda consiguiendo individuos casi perfectos, esto se refiere a que el nivel de error se aproxima a 0. *SAPIFE*³ también logra conseguir ajustes muy buenos, y vuelve a repetir la ocurrencias de las suposiciones anteriores, donde *SAPIFE*³ se aproxima bastante al mejor método en estas circunstancias (Evolutivo) y consigue mejor resultado que *S²F²M*. Por otra parte, se nota en los últi-

mos periodos de tiempo, que el incendio ha modificado en cierta medida su comportamiento y los métodos cambian de su posición en el ranking de efectividad, excepto *SAPIFE*³. En otras palabras, el método que fue mejor en un periodo regular del incendio, pasa a ser el peor en un periodo un tanto irregular. Sin embargo, *SAPIFE*³, ha conseguido un equilibrio manteniéndose siempre como el método más regular. Este resultado, al igual que los anteriores permite comprobar con cierta certeza, que la combinación de los dos métodos ha logrado un equilibrio de las fortalezas y debilidades de ambos en el ajuste.

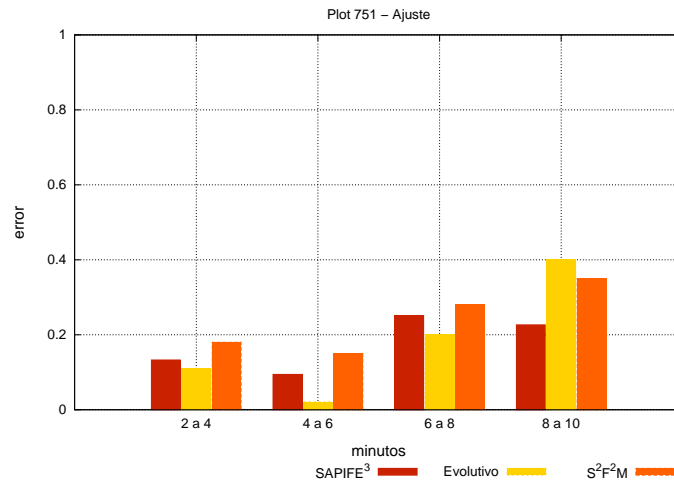


Figura 5.15: Plot 751, errores obtenidos en la etapa de ajuste

En la figura 5.16, se observan los resultados conseguidos por los métodos en la etapa de predicción, esto permite afirmar con mayor fuerza, que en un incendio con terreno homogéneo y condiciones medio-ambientales estables, el ajuste determina la calidad de la predicción. Si se pone especial atención, a los dos primeros pasos de tiempo, se observa que el Evolutivo ha conseguido un nivel de error muy bajo, especialmente en el periodo de tiempo de 4 a 6, donde la predicción es casi perfecta y esto se relaciona con los resultados obtenidos en la etapa de ajuste. Al llegar el tercer paso de tiempo, se puede observar que el error tiene una tendencia a crecer de nuevo, esto podría ser, debido a que las condiciones climatológicas empiezan a variar, lo que se ve reflejado en el mapa de propagación de la figura 5.14, que en el último periodo de tiempo, la propagación del incendio se frena considerablemente. En esta situación de cambio, se observa que *S*²*F*²*M* y *SAPIFE*³ consiguen resultados de error más bajos que el Evolutivo y de hecho, esto es razonable, puesto que en los casos anteriores, se demostró que estos métodos están mejor diseñados para soportar cambios de contexto en las condiciones del clima,

entre el periodo de ajuste y predicción.

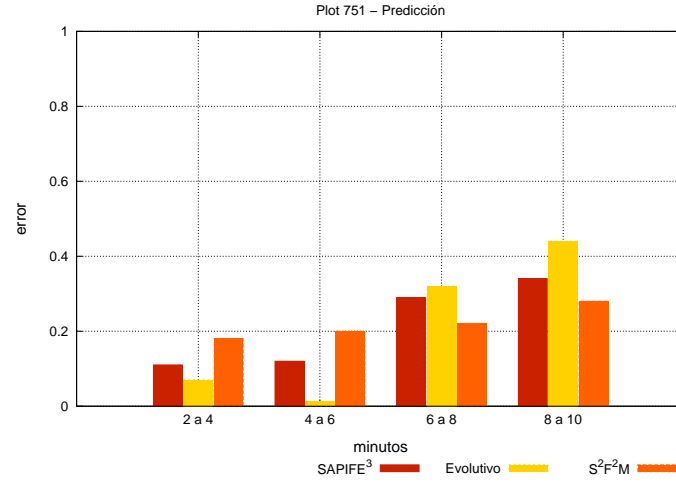


Figura 5.16: Plot 751, errores obtenidos en la etapa de predicción

El estudio de los casos de mapas reales, sirvió principalmente para demostrar que *SAPIFE*³, consigue con la integración del Evolutivo y *S*²*F*²*M*, utilizar las cualidades de uno para suplir las carencias del otro, y viceversa. De este modo, se obtiene un método más equilibrado, donde la adaptación en condiciones ambientales, en constante cambio, es una necesidad de vital importancia para elaborar la predicción. También, con estos experimentos se puede tener una idea más acabada, de las situaciones que deben prever los simuladores de fuego, y del porque, es tan complejo predecir este tipo de fenómeno.

Capítulo 6

Conclusiones y Trabajo Futuro

6.1. Consideraciones

Los incendios forestales siempre han sido un agente de cambio en la naturaleza, esto ha hecho que muchas especies se hayan adaptado a la coexistencia con el fuego haciéndose resistentes o migrando a zonas más extensas. Pero, el verdadero problema está cuando las causas no son por factores naturales y la frecuencia de ocurrencias de estos fenómenos aumenta, esto produce que las especies no puedan recuperarse a tiempo para el siguiente incendio.

Las consecuencias ecológicas, sociales y económicas presentan cifras alarmantes, es por ello, que muchos recursos e investigaciones se encuentran avocados a la lucha contra los incendios forestales, que se considera, ya no depende sólo de factores naturales. Pero para encarar este tipo de problemas se debe conocer sus causas, comportamientos y consecuencias.

En este trabajo se ha estudiado este fenómeno e identificado dónde se puede realizar una aportación que ayude a prevenir las consecuencias catastróficas, apoyados en la ciencia y los avances tecnológicos en el área de la supercomputación, para lograr de alguna forma mitigar los daños que provocan los incendios forestales.

Se ha prestado principal atención al comportamiento del incendio, esto permite grandes ventajas en la coordinación de tareas para la prevención, gestión del manejo de operaciones, de lucha y rescate. Entonces, si se conoce cómo es el comportamiento también se puede predecir y de esta forma, ser mucho más eficientes en las decisiones y evitar pérdidas de vidas humanas, ecológicas y materiales.

También, se ha dado especial énfasis a cómo la tecnología puede ayudar en la lucha contra los incendios forestales, este fenómeno físico requiere de todos los recursos disponibles: humanos, económicos y tecnológicos. El estudio de este problema, es una tarea que requiere de la intervención de distintas áreas de la ciencia, es por ello, que este trabajo se enmarca en una nueva forma de hacer ciencia, como lo es la *Ciencia Computacional*.

Al iniciar la memoria, los objetivos estaban claros, y bien establecidos pero el desafío estaba en cómo se podría hacer. Para ello, fue necesario conocer el fenómeno, las causas, el comportamiento, las soluciones y lo que sirvió de base a esta memoria, las limitaciones de las soluciones existentes.

El trabajo esta fuertemente ligado a la simulación y los modelos matemáticos con que trabajan, y en este punto para fundamentar aún mejor esta idea se menciona una frase escrita por John von Neumann (1903-1957):

Las ciencias no tratan de explicar, incluso apenas tratan de interpretar, construyen modelos principalmente. Por modelo, se entiende una construcción matemática que, con la adición de ciertas interpretaciones verbales, describe los fenómenos observados. La justificación de tal construcción matemática es sólo y precisamente que se espera que funcione.

Por lo tanto, la tarea no consiste en construir modelos exactos, ni simuladores avanzados y detallistas, sino simplemente hacer que éstos funcionen.

6.2. Conclusiones

Para realizar este trabajo, primero se debía comprender donde estaba situado, es decir, en que área de la ciencia específicamente, y luego del análisis de campo, se concluyó que el entorno de esta investigación está definido como la “*Ciencia Computacional*”. Los tres pasos que llevan a este razonamiento se puede representar imaginariamente como los vértices de un triángulo, al que se podría llamar “Triángulo de Fuego”:

- **Ciencia Aplicada:** El primer paso fue comprender el funcionamiento del fuego, y para ello era necesario un modelo físico para representar de manera conceptual este proceso (fenómeno), con el fin de analizar su naturaleza, desarrollar o comprobar hipótesis o supuestos y permitir una mejor comprensión de este fenómeno real.
- **Matemática:** El siguiente paso fue la necesidad de tener un lenguaje más preciso que permita analizar este fenómeno numéricamente, y esto sólo se podía hacer por medio de un modelo matemático. Se puede decir que este modelo es una traducción de la realidad física para poder aplicar los instrumentos y técnicas de las teorías matemáticas y estudiar el comportamiento de sistemas complejos, y posteriormente hacer el camino inverso para traducir los resultados numéricos a la realidad física. Y esto llevó a la confirmación que el modelo matemático es sólo una simplificación de la realidad y es prácticamente imposible modelar todos los detalles de este fenómeno natural.
- **Informática:** Por último, se observó que traducir este modelo a un lenguaje de máquina, era demasiado complejo para un computador convencional y exigía un alto poder de cómputo, entonces, fue necesario

recurrir a las ciencias de la computación, de tal forma a conseguir mejorar los resultados de las simulaciones con cómputos intensivos de datos, utilizando técnicas como la paralelización.

El estudio del arte en incendios forestales, llevó a entender en gran medida el funcionamiento de los distintos simuladores de fuego, los modelos con que trabajan, sus implementaciones y lo más interesante para esta memoria, las limitaciones de cada uno de ellos. Aquí se observó un punto de convergencia, si bien cada simulador estaba implementado con un modelo distinto, diferente lenguaje de programación, con menor o mayor precisión, todos presentaban el mismo problema, eran dependientes de la medición de las variables que se utilizaban como entrada, éstas contenían un alto grado de incertidumbre e imprecisión. El otro factor era el comportamiento de las variables dinámicas, muchos de los parámetros de gran influencia eran sumamente difíciles de medir y el grado de precisión que necesita la propia física del modelo implementado en el simulador, es prácticamente imposible de determinar.

El trabajo se enfocó en las metodologías guiadas por los datos, de solución única (Evolutivo) y solución múltiple (S^2F^2M), que comprenden técnicas muy prometedoras, donde el simulador es una herramienta más dentro del modelo, y proponen una perspectiva diferente. Este paradigma, cambia el sentido de la predicción tradicional, introduciendo una nueva etapa antes de la predicción, de tal manera a ajustar el simulador interviniendo sobre los datos con los cuales operan.

Con estas metodologías se consiguieron mejorar muchos aspectos de la predicción tradicional, pero concientes de que todo puede ser mejorado se analizó cuáles eran los puntos fuertes y débiles de las dos metodologías. Este estudio obligó a una mayor comprensión del funcionamiento interno de ambos métodos, razón por la que se dedicó gran parte del tiempo y esfuerzo en comprender las distintas metodologías e incluso proponer algunas mejoras sobre las mismas.

A raíz de estos estudios, se ha desarrollado una propuesta para mejorar la predicción, basado en las fortalezas y debilidades de las metodologías de solución única y solución múltiple, para ello, se ha identificado el problema principal que se desea atacar y se han definido las causas y las posibles soluciones. El evolutivo es un método que puede llegar a ser muy eficiente pero en algunos casos tiene la dificultad para converger en una solución aceptable y es muy dependiente de la configuración que pueda tener, además de ser algo susceptible al cambio de las condiciones ambientales del incendio. Sin embargo, el método estadístico es menos susceptible a estos cambios con algunas dificultades para cierto tipo de incendios (frentes de fuego elípticos con una excentricidad muy angosta) y un alto requerimiento de cómputo por la gran cantidad de simulaciones que necesita realizar.

El método evolutivo tiene como semilla un algoritmo genético *AG*, que es una técnica de optimización de parámetros muy potente que podría ser

utilizado como una herramienta más de S^2F^2M , que es un método basado en fundamentos estadísticos, en donde internamente utiliza la frecuencias de sucesos barriendo una gran parte del espacio de las posibilidades que se podría dar en un incendio. Esto permite adelantarse en el tiempo y predecir sin la necesidad de conocer los datos exactos del clima.

Con estas herramientas, la idea principal es utilizar el AG para optimizar las entradas de S^2F^2M y reducir el número de simulaciones y con ello, el tiempo para predecir, de esta forma se busca dotar de mayor equilibrio a todo el sistema consiguiendo buenos parámetros y construyendo predicciones menos sensibles al cambio de las condiciones del clima.

Entonces, para integrar estas dos metodologías, se debió zanjar diferencias muy sensibles, de tal modo que esta propuesta conserve las características fundamentales de cada una, definiendo sus principios básicos, cuyo desarrollo dio como resultado el Sistema Adaptativo de Predicción de Incendios Forestales basados en Estrategias Estadístico-Evolutivas ($SAPIFE^3$). Este nuevo sistema también exige una gran capacidad de cómputo, por lo que la aplicación fue desarrollada para un entorno paralelo. Es posible que la implementación de este nuevo método no esté basado en un esquema de paralelización óptimo. Pero esta implementación puede ser mejorada.

Con el fin de comprobar si esta propuesta era viable se elaboraron una serie de experimentos de los cuales se han llegado a las siguientes conclusiones:

- En los experimentos, se puede notar que en un entorno uniforme y condiciones estables el ajuste determina la calidad en la predicción, pero en un periodo de inestabilidad el error en la predicción se dispara y ya no depende del ajuste. Esto es razonable, ya que el ajuste esta basado en un estado de condiciones diferentes al de la predicción. En este caso se podría reducir el rango de tiempo de los periodos.
- En situaciones donde se presenta un cambio brusco de las condiciones, entre la etapa de ajuste y predicción, el evolutivo tiene un salto en el nivel de error en la predicción superior a $SAPIFE^3$ que consigue un menor efecto negativo en estas circunstancias, por lo tanto, logra adaptarse más rápido.
- En situaciones donde los incendios tienen un comportamiento de un frente de fuego elíptico con una excentricidad muy angosta, el método S^2F^2M tiene gran dificultad para ajustarse y gracias a la integración este problema es mejorado considerablemente.
- En un incendio con un frente de fuego lineal, esta nueva propuesta reduce el número de simulaciones y el nivel de error es equiparable con S^2F^2M que exige muchas más simulaciones y tiempo para elaborar la predicción que $SAPIFE^3$.

- En los casos donde el método evolutivo, que normalmente tiene un mejor ajuste con errores mas bajos, *SAPIFE*³ logra acercarse bastante.
- Se observó que *SAPIFE*³ también es dependiente de la configuración del algoritmo evolutivo, por lo que primero fue necesario calibrar estos parámetros, de tal manera a obtener mejores resultados y este aspecto se debe tener en cuenta para utilizar alguna estrategia que rompa esta dependencia.
- En todos los experimentos se considera al entorno de simulación uniforme, pero esto en la realidad rara vez ocurre, el terreno es casi siempre heterogéneo, es por ello, que los resultados en quemas reales el nivel de error es superior que en los casos de incendios sintéticos, además de que las condiciones climatológicas cambian.
- En los resultados con los mapas de quemas reales, se observó que *SAPIFE*³ consigue un buen equilibrio y algunos puntos producen un mejor ajuste y predicción que el resto de los métodos, y los pasos de simulación que necesita para recuperarse es menor.

Se considera necesario un mayor número de casos para realizar experimentos con un análisis más exhaustivo para su validación definitiva, pero dentro del marco de los objetivos propuestos para cada caso se alcanzó un nivel satisfactorio en los resultados.

6.3. Trabajo Futuro

Es necesario realizar un estudio sobre los índices de impacto de cada variable, y establecer una jerarquía de peso. Con esto se puede determinar cuáles son los parámetros de mayor influencia, que servirán para elaborar librerías de datos y tratarlos con algunas técnicas de Data-Meaning.

En la primera fase del trabajo se logró resultados muy alentadores, se consiguió probar que este nuevo sistema tiene una adaptación más rápida a condiciones cambiantes del clima. Una tarea a realizar es el estudio de como hacer un mejor ajuste y elaborar mapas de probabilidades para distintas situaciones.

Es necesario realizar un estudio más exhaustivo de la relación % de error en la predicción frente al tiempo invertido (número de simulaciones), para valorar la aplicación del método en un entorno de tiempo real.

Sería importante realizar la prueba del método con otro simulador como núcleo y de esta manera comprobar que la metodología es independiente del simulador con que se trabaje. Para este trabajo se ha utilizado el simulador fireLib, que gracias a su simplicidad ha contribuido a conocer el

funcionamiento interno de un simulador de fuego. Ahora, la tarea es aplicar este conocimiento incorporando otros tipos de simuladores y probar que esta metodología puede ser utilizada para potenciar cualquier otro simulador.

Se debe elaborar alguna estrategia donde el algoritmo evolutivo, en caso de acercarse a zonas que no pueda converger en una buena solución, pueda realizar saltos en el espacio de búsqueda, explorando de esta manera, otras posibilidades que mejoren la calidad de los individuos.

Hasta este momento se ha considerado superficies totalmente homogéneas, el desafío siguiente está en construir el método considerando superficies heterogéneas. Una idea sería particionando el terreno en varias superficies y que cada una represente una superficie homogénea distinta, de esta forma también, se puede considerar que cuando el incendio adquiere cierta extensión las condiciones ya no son las mismas en todos los puntos del incendio. Por lo tanto, cada superficie homogénea puede tener su propia condición climatalógica y generar su propio mapa de propagación, para luego integrarlo en un mapa general de la propagación. Por supuesto, que esto va a requerir una gran capacidad de cómputo y presenta una problemática bastante compleja en cuanto a las fronteras de integración que debe ser analizada.

Bibliografía

- [1] G. Bianchini “Sistema de Ayuda a la Decisión Para la Gestión de Incendios Forestales,” master’s thesis, Dept. Computer Architecture and Operating Systems, Universitat Autònoma de Barcelona, 2004.
- [2] P. J. Denning, “Is Computer Science Science?,” *Commun. ACM*, Apr. 2005, pp. 27-31
- [3] Tapia R. A., Lanius C., Mc Zeal C. M., Parks T. A. “Computational Science: Tools for a Changing World,” Rice University, 3 Jul. 2003; <http://ceee.rice.edu/Books/CS/index.html>
- [4] D. J. Mostaccio, “Simulación de Altas Prestaciones para Modelos Orientados al Individuo,” doctoral dissertation, Dept. Computer Architecture and Operating Systems, Universitat Autònoma de Barcelona, 2007.
- [5] T. H. Naylor, J. L. Balintfy, D. S. Burdick and K. CHU, *Técnicas de Simulación en Computadoras*, Limusa-Wiley, pp. 15-35, 1982.
- [6] A. Rosenblueth and N. Wiener, *The Role of Model in Science*, Aldine Publishing Co., 1963, pp. 316-321.
- [7] M. A. Law ,W. Kelton, *Simulation Modeling & Analysis*, McGraw-Hill, pp. 1-109, 1991.
- [8] J. Glasa and L. Halada, “Application of envelope theory for 2d fire front evolution,”*Proc. 5th Int’l Conf. Forest Fire Research*, CD-ROM, 2006.
- [9] D. Caballero, “Taxicab geometry: Some problems and solutions for square grid-based fire spread simulation,”*Proc. 5th Int’l Conf. Forest Fire Research*, CD-ROM, 2006.
- [10] R. C. Rothermel, *How to predict the spread and intensity of forest and range fires*, tech. report INT 143, Agriculture Dept. Intermountain Forest and Range Experiment Station Ogden,U.S. Forest service, 1983.
- [11] A. Grama ,A. Gupta, G. Karypis, V. Kumar, *Introduction to Parallel Computing*, Addison Wesley, pp. 12-76, 2003.

- [12] P. Johnston, G. Milne, "A heat transfer simulation model for wildfire spread," *Proc. 5th Int'l Conf. Forest Fire Research*, CD-ROM, 2006.
- [13] J. Coleman and A. Sulliman, "SiroFire: the CSIRO bushfire spread simulator," *Proc Inst Forest Aust 16th Biennial Conf.*, 1995, pp. 309-319.
- [14] M. A. Finney, *FARSITE: Fire area simulator - model development and evaluation*, tech. report RMRS-RP-4, Agriculture Dept. Rocky Mountain Research Station Ogden, U.S. Forest service, 1998.
- [15] Tymstra, "The Canadian Wildland Fire Growth Model," Feb, 2007; <http://www.firegrowthmodel.com/software.cfm>.
- [16] I. Knight. and J. Coleman, " A fire perimeter expansion algorithm based on Huygen's wavelet propagation," *International Journal of Wildland Fire*, vol. 3 no. 2, pp. 73-84, 1993.
- [17] T. L. Clark, M.A. Jenkins, J. Coen and D. Packham, " A Coupled Atmosphere-Fire Model: Convective Feedback on Fire-Line Dynamics," *Journal of Applied Meteorology*, vol. 35 no. 6, pp. 875-901, 1996.
- [18] R. Linn, J. Reisner, J. J. Colman and J. Winterkamp, " Studying wildfire behavior using FIRETEC," *International Journal of Wildland Fire*, vol. 11 no. 3, pp. 233-246, 2002.
- [19] F. J. Serón, D. Gutiérrez, J. Magallón, L. Ferragut, L. and M. I. Asensio, " The Evolution of a WILDLAND Forest FIRE FRONT," *The Visual Computer: International Journal of Computer Graphics*, vol. 21 no. 3, pp. 152-169, 2005.
- [20] S. G. Berjak and J. W. Hearne, " An improved cellular automaton model for simulating fire in a spatially heterogeneous Savanna system," *Ecological Modelling*, vol. 148 no. 2, pp. 133-151, 2002.
- [21] A. Dunn and G. Milne, "Modelling Wildfire Dynamics via Interacting Automata," *Proc. 6th International Conf. on Cellular Automata for Research and Industry.*, Springer, 2004, pp. 395-404.
- [22] E. H. Anderson, *Predicting wind-driven wildland fire size and shape*, tech. report INT-305, Agriculture Dept. Intermountain Forest and Range Experiment Station Ogden, U.S. Forest service, 1983.
- [23] D. B. Collins, "FireLib User Manual & Technical Reference," Nov, 2006; <http://www.fire.org/>
- [24] P. L. Andrews, *BEHAVE: Fire Behavior prediction and modeling systems Burn subsystem, part 1*, tech. report INT-194, Intermountain Forest and Range Experiment Station, U.S. Forest service, 1986.

- [25] M. M. Denham, “Predicción de Incendios Forestales basada en Algoritmos Evolutivos Guiados por los Datos,” master’s thesis, Dept. Computer Architecture and Operating Systems, Universitat Autònoma de Barcelona, 2007.
- [26] J.R. Koza, *Genetic Programming. On the Programming of Computers by Means of Natural Selection*, The MIT Press, pp. 819, 1992.
- [27] G. Bianchini, “Wildland Fire Prediction based on Statistical Analysis of Multiple Solutions,” doctoral dissertation, Dept. Computer Architecture and Operating Systems, Universitat Autònoma de Barcelona, 2006.
- [28] Gnuplot homepage, “gnuplot An Interactive Plotting Program Version 4.2,” Mar, 2007; <http://www.gnuplot.info/index.html>.
- [29] E. Heymann, “Effective Resource Management for Master-Worker Applications in Opportunistic Environment,” doctoral dissertation, Dept. Computer Architecture and Operating Systems, Universitat Autònoma de Barcelona, 2001.
- [30] SPREAD homepage, “Project Abstract,” Jul, 2007; <http://www.algosystems.gr/spread/index.html>